UNIVERSITY OF BIRMINGHAM

Photochromic Control of Nucleic Acid Systems through Anthracene Photodimerisation

Jack Nathaniel Manchester

A thesis submitted to the University of Birmingham for the degree of Doctor of Philosophy

School of Chemistry

College of Engineering and Physical Sciences

The University of Birmingham

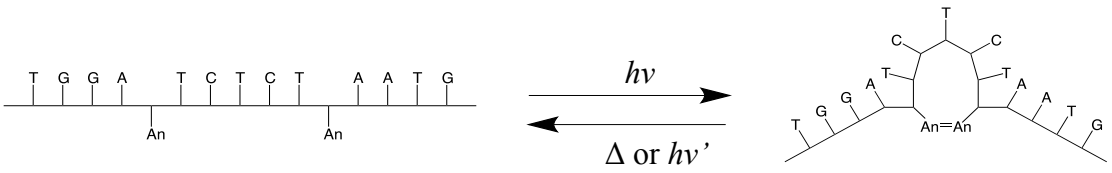
September 2013

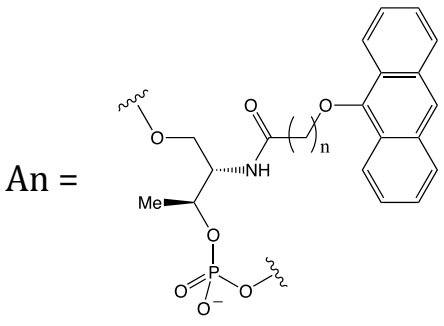
This thesis is dedicated to Deborah Mary Hogan

Abstract

The focus of this project has involved the incorporation of non-nucleosidic anthracene phosphoramidites into oligonucleotide sequences for the purpose of studying anthracene photodimerisation and its effect on DNA structure and function. The primary aim of the research was to use this technique as a method of disrupting hybridisation between complementary oligonucleotides, but has been subsequently expanded to study the effect on other forms of nucleic acid structure.

By introducing two anthracene tags into the same oligonucleotide sequence it is possible to initiate a reversible, light-induced reaction involving the intramolecular dimerisation of the appended anthracene groups. The resulting change in structure, combined with a change in π-stacking interactions within the oligonucleotide, alters the binding towards the complementary sequence. This effect was studied through a number of techniques including variable temperature UV spectroscopy, CD and fluorescence spectroscopy and gel electrophoresis.





The system possesses a dual-mode gated behaviour, in which the process of photodimerisation can be supressed by preforming the duplex, which limits the proximity of anthracene due to a number of factors. However, attempts were made to achieve in-*situ* 'photo-release' of a bound complementary sequence by the inclusion of short, complementary overhang domains, which can initiate the formation of stem-loop structures.

Modification of the G-Quadruplex nucleic acid structure was also studied and through a number of different strategies the effect of photodimerisation was used to alter the behaviour of the folded oligonucleotide. The dependence of the G-Quadruplex on the presence of a templating metal cation adds an additional input to the method and has been shown to alter the photochemistry of certain sequences.

Finally the intermolecular photodimerisation of anthracene modified oligonucleotides was studied in order to achieve photo-controlled interstrand cross-linking. The dimerisation of complementary sequences was found to greatly stabilise the duplex towards a number of factors including changes in temperature and ionic strength.

Acknowledgements

There are a great number of people who have contributed towards my research within this project and without their help this thesis would not be possible, for this I am eternally grateful.

Firstly I would like to thank Dr John Zhao for training me in all aspects of oligonucleotide synthesis, purification and analysis, as well as guidance in many other techniques. For helping to set up the photochemical studies I would like to thank Dr Luciana Giordano, and to Dr Dario Bassani of the University of Bordeaux for his help in analysing the results and perfecting the technique. For contributing materials that have greatly aided the efforts of this project as well as providing his advice I would also like to thank Dr Jean-Louis Duprey.

At the University of Birmingham I would like to thank Dr Neil Spencer for assisting with NMR analysis and Peter Ashton for mass spectrometry. A big thank you to the now retired Graham Burns for teaching me all there is to know about HPLC and for his daily banter that I now so sorely miss. Another big thank you to Dr Chi Tsang who more than excelled in filling Graham's shoes on his retirement. I would also like to acknowledge the support of Erasmus student Rafal Ramzi for assisting me in the synthesis of additional anthracene monomers and performing other techniques, which have supported this project. I would also like to thank Dr Anna Peacock, my second supervisor, for her support and advice throughout this project.

Within the Tucker group I would like to thank all members, past and present, for their kind support over the years, they truly are an exceptional group of people and I am proud to have worked alongside them. In particular I would like to thank my partners in crime and fellow musketeers Antoine Sullastrau and Dr Andrea Mulas for all the good times we

have shared together over the years. I am also indebted to Gemma Bullen and Rosemary Bamford for their superb proofreading skills and for all those funny moments that still put a smile on my face. Thanks also go to Dr Huy Nguyen, Dr Giorgio Mirri, Michel Boissonnet, Pete Thornton and James Carr-Smith, to all of them I wish the brightest of futures.

Over the past few years I would have been unable to achieve anything if it were not for the help and encouragement from my friends and family and it is with sincere gratitude that I acknowledge them here. In particular my thanks go to Alex, Marie, Pam, Carlotta and Kevin, although we go our separate ways our friendships will never end.

Finally I would like to thank Dr James Tucker for not only giving me this fantastic opportunity but for his guidance and support, which have been invaluable over the years.

Table of Contents

Abbreviations

General Abbreviations:

a.q.	Aqueous
CH_3CN	Acetonitrile
AFM	Atomic Force Microscopy
CL	Cross-Link
CPG	Controlled Pore Glass
D_1 / HT	Anthracene Head to Tail Photodimer
D_2 / HH	Anthracene Head to Head Photodimer
DCM	Dichloromethane
DIPEA	*N,N*---diisopropylethylamine
DMAP	*N,N*---dimethylaminopyridine
DMF	*N,N*---dimethylformamide
DMTr	Dimethoxytrityl
DNA	Deoxyribonucleic Acid
ETT	2-Ethylthiotetrazole
FRET	Förster Resonance Energy Transfer
GE	Gel Electrophoresis
HBTU	O---Benzotriazole---*N,N,N',N'*---tetramethyl---uronium---hexafluoro---phosphate
HPLC	High Performance Liquid Chromatography
ICL	Interstrand Cross-Link
ITC	Isothermal Titration Calorimetry
Me	Methyl
µL	Microliter
µM	Micromole
mM	Millimole
nm	Nanometer
NAP	Nucleic Acid Purification
NMI	*N*-methylimidizole
PDB	Protein Data Bank
Py	Pyridine
SL	Stem Loop
SM	Starting Material
SNP	Single Nucleotide Polymorphism
TB	Tris-Borate buffer
TBA	Thrombin Binding Aptamer
TBE	Tris-Borate-EDTA
TCA	Trichloroacetic acid
TEA	Triethylamine
TEAA	Triethylammonium Acetate
THF	Tetrahydrofuran

T_m	Melting Temperature

Spectroscopic / Photochromism Terms:

CD	Circular Dichroism
$h\nu$	Light Energy
UV	Ultra Violet Light
Vis	Visible Light
VT-UV	Variable Temperature Ultra Violet Spectroscopy
Δ	Heat Energy
ΔA	Change in Absorbance
ε	Extinction Coefficient
λ	Wavelength

DNA Terms:

CPG	Controlled-Pore Glass
A	Adenine
C	Cytosine
G	Guanine
T	Thymine

Chapter 1: Introduction

1.1 Supramolecular Chemistry

1.1.1 Introduction to Supramolecular Chemistry

In the words of Jean-Marie Lehn, supramolecular chemistry is, in a way, a form of "molecular sociology", a study of the various interactions between molecules and how it can lead to the development of complex ordered systems, constructed from a number of discrete subunits.[1] He, along with scientists Donald J. Cram and Charles J. Pedersen were awarded the Nobel Prize for Chemistry in 1987 for their work on the synthesis of molecules that mimic important biological processes. It was this work that laid the foundations for what has now become a progressively diverse new branch of chemistry.[2]

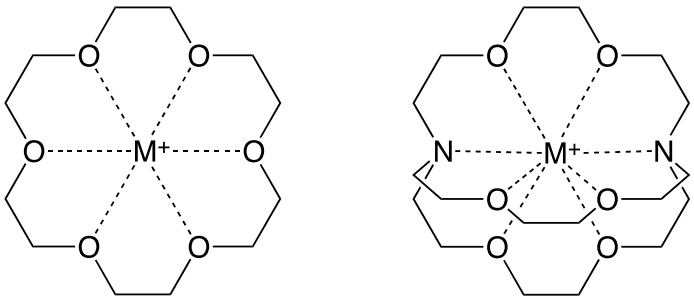


Figure 1. Examples of crown and cryptand structures binding a metal cation (M^+) guest through electrostatic interactions.

It began with the study of crown and cryptand structures (Figure 1), which can bind to cationic guests, such as sodium or potassium through electrostatic interactions and can discriminate between different ions through a size-fit approach.[3] This early example demonstrated how it was possible to design structures to exploit multiple non-covalent interactions, which individually are quite weak, but together can cooperate to form strong complexes.

Since that time supramolecular structures have become ever more complex in design, taking advantage of a wide variety of intermolecular interactions such as hydrogen bonding and π-π stacking, and the field itself has expanded to cover a wide number of disciplines from coordination to bioorganic chemistry.[4] At its core are a number of fundamental concepts, which are vital to the design of any supramolecular system, as described in the following sections.

1.1.2 Molecular Recognition

For a supramolecular system to take advantage of multiple intermolecular interactions its substituent components must exhibit some degree of complementarity.[5] This is important in a biological sense; the Dutch chemist Emil Fisher proposed in 1894 that an enzyme and substrate fit together "like lock and key", to describe how the active site within an enzyme is perfectly shaped to bind a specific substrate in order to carry out a biological process.[6] Chemists have often tried to replicate this property through the design of molecular architectures maximising interactions towards a specific guest, making it selective.

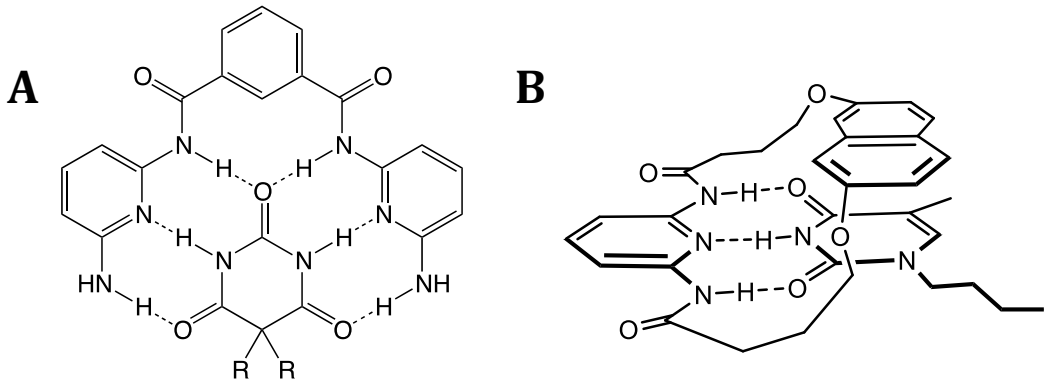


Figure 2. Examples of supramolecular receptors binding to a barbiturate (A) and a thymine derivative (B) through multiple non-covalent interactions.[8]

Hamilton *et al.* has been successful in producing a number of receptors, each one having a structure that has been tailored to recognise the specific shape of a target species. Incorporating specific functional groups within these motifs, which further complement the features of the target molecule, lead to the formation of the non-covalent interactions essential for the development of a supramolecular complex. As an example barbiturates can be bound through an array of six hydrogen bonds (Figure 2A),[7] whilst a thymine nucleobase can be bound through a combination of hydrogen bonding and π-stacking interactions (Figure 2B).[8]

This area of research has been termed as 'Host-Guest' chemistry, and has often been used for sensing purposes. The presence of a guest species can be detected by incorporation of reporting groups within the design of the host structure. Thus, the binding of the guest induces a change that is detected through a variation in the response given by the reporter group. A good example of this type of work is that of De Silva *et al.* who developed anthracene functionalised crown ethers for the sensing of alkali metal ions (Figure 3). In the unbound state fluorescence from the anthracene modification is quenched by electron transfer

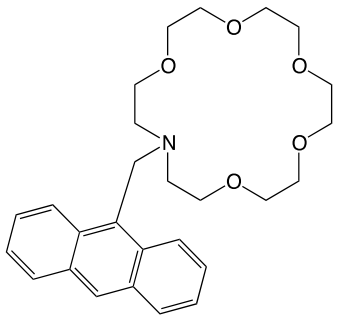


Figure 3. An anthracene functionalised crown ether for the fluorescent sensing of alkali metal ions.

from the lone pair on the nitrogen. But upon binding of a metal cation electron transfer no longer occurs, with the lone pair coordinated to the cation, therefore leading to a significant increase in the fluorescent quantum yield and allowing the presence of the cation to be detected.[9]

1.1.3 Self Assembly

Molecular self-assembly is the process by which molecules assemble spontaneously, without the guidance from an external source, to create organised structures from disordered systems.[10] The process is driven by the formation of non-covalent interactions, since self-assembly is entropically unfavourable. This is explicitly linked to molecular recognition, mentioned earlier, where the design of the final structure is encoded for in the shape and properties of the subunits that assemble to form the complex. Self-assembly can occur intramolecularly, for example when a polypeptide folds into a functional protein structure, or intermolecularly, such as in the previous examples (Figure 2).

1.2 DNA

DNA is a classic example of a supramolecular system, which has had the benefit of millions of years of evolution to perfect its design. It is a polymer consisting of an arrangement of monomers known as nucleotides, which in turn are comprised of a heterocyclic base, a deoxyribose sugar and a phosphate residue (Figure 4A).[11]

The nucleotides within each strand are linked to each other from the 3' carbon of one deoxyribose to the 5' carbon of the next through a phosphate group; DNA is therefore described as having a direction, moving from the 5' to 3' end. It is this direction that is used to describe the order of bases when listing DNA sequences.

There are four main types of heterocyclic base within DNA, two purines adenine (A) & guanine (G) and two pyrimidines cytosine (C) & thymine (T). The complementary arrangement of hydrogen bond donor/acceptor groups within these structures leads to the formation of Watson-Crick pairs (Figure 4B), named after the two scientists who discovered the structure of DNA in 1953.[12] This form of molecular recognition makes each nucleobase selective for its Watson-Crick pair, with the G/C pair able to form three hydrogen bonds, whilst the A/T pair forms two.

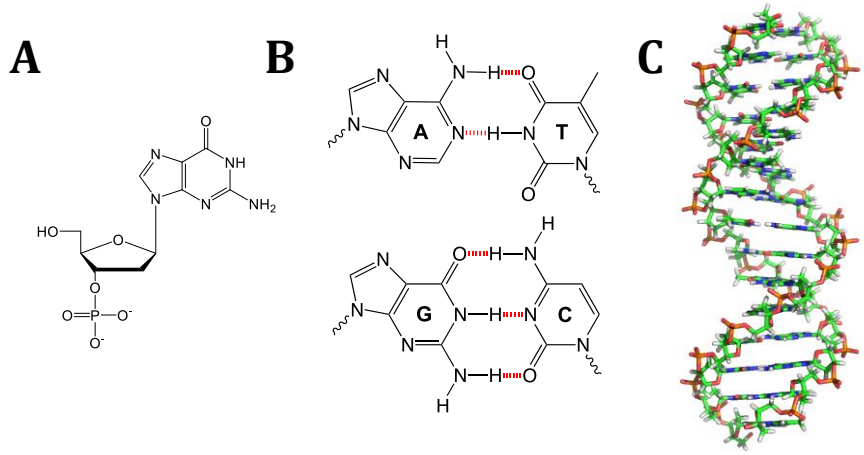


Figure 4. (A) Structure of a nucleotide, in this case guanine, (B) Hydrogen bonding between Watson-Crick base pairs, (C) The double helix structure of duplex DNA (B-type).

In solution, two single stranded segments of DNA oriented anti-parallel to each other will spontaneously self-assemble into a duplex structure through a process known as hybridisation, where each strand is linked through a network of hydrogen bonds within the Watson-Crick pairs. This will only occur if the sequence of bases within each strand is complementary to the other, making

hybridisation a highly selective process. The stability of a duplex is mainly determined by the compatibility of the two sequences and through maximising the number of effective hydrogen bonds they form upon hybridisation. However hydrogen bonding is not the only factor that contributes to the formation and stability of a DNA duplex. The end result of DNA hybridisation is the formation of a double helical structure (Figure 4C). The structure maximises the overlap of the heterocyclic nucleobases, allowing for a greater number of electrostatic π-stacking interactions between the bases. In addition to this the structure increases the distance between the phosphate groups within the backbone, therefore minimising electrostatic repulsions, however these repulsions are mainly overcome by the presence of counter ions in solution. These combined factors provide the duplex structure a greater stability over its single stranded components.[13]

There are three main types of DNA secondary structure that duplexes can adopt (A, B and Z) and their formation mainly depends on both sequence composition and relative humidity. A and B-DNA are both right handed helical structures, with A-DNA having a slightly tighter twist under lower humidity conditions and Z-DNA is a left handed helix that forms mainly in alternating purine/pyrimidine sequences (e.g. CGCGCG). Each structure has different levels of compactness with A, B and Z forms having 11, 10.5 and 12 base pairs per helical turn respectively. Many studies have shown that the B-DNA type structure forms predominantly in cellular conditions.[14]

1.3 DNA Nanotechnology

The primary role of DNA is to store the genetic information required to create life, programmed like a computer language into the structure through the specific arrangement of bases within each sequence. But since the discovery of its structure and the realisation of its unique properties, paired with recent developments in oligonucleotide synthesis, research into DNA has begun to branch out into the realm of nanotechnology. [14]

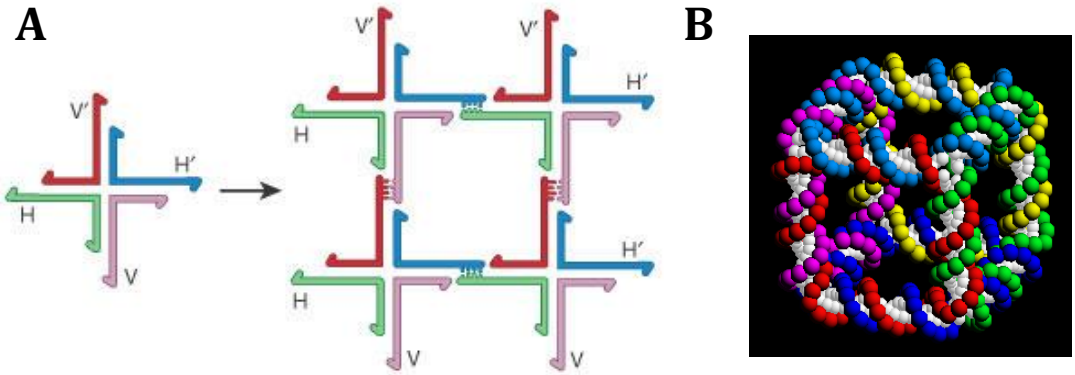


Figure 5. (A) The development of fixed, four way DNA junctions into tiled nano-arrays. (B) A three dimensional DNA assembly consisting of six interlinking strands to form a DNA cube. Image taken from ref. 16.

It is now possible to create complex DNA-based nanostructures, which are fabricated with precision due to DNA's unique self-assembling properties, governed by a strict set of base pairing rules. This research began in the early eighties when Nadrain 'Ned' Seeman and co-workers looked at natural branched DNA junctions, known as Holliday junctions, for inspiration in the creation of larger, two-dimensional frameworks.[16] By rationally designing the order of bases within a set of sequences so that they self-assemble into a fixed four-way junction, followed by the hybridisation of short, single-stranded overhangs known as 'sticky ends', it was possible to create an array of repeating tile structures (Figure 5A).[17] After further research Seeman was able to move into a

third dimension in order to generate a DNA cube, using similar strategies as before to interlink six cyclic oligonucleotide sequences, with each one forming a single side of the shape (Figure 5B).[18]

The main drive of this research has been to organize materials at the nano-scale level into designed patterns in a 'bottom-up' approach, thus gaining control over the interactions between certain nanoparticles.[17,19] The collective properties of these nanoparticles, such as electron transport or optical coupling are dependent on their relative arrangement, therefore utilising the programmable features of DNA has led to the development of functional devices. Kiehl *et al.* demonstrated how it is possible to use a DNA scaffold, similar to those developed by Seeman, to arrange two differently sized gold nanoparticles through selective hybridization into precise rows (Figure 6A).[20,21] The gold nanoparticles were functionalised with short single stranded DNA molecules designed to complement specific binding sites within the DNA scaffold, utilising Watson-Crick pairing to dictate the spatial arrangement of the particles with respect to each other. The idea

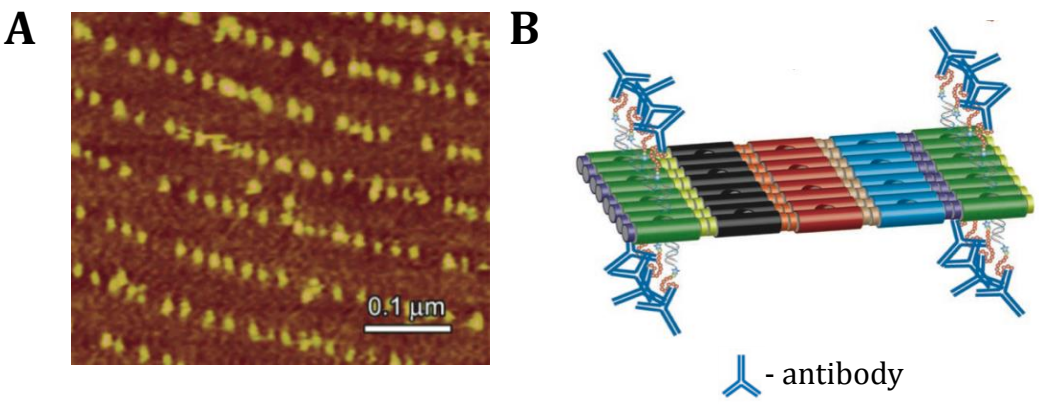


Figure 6. (A) Atomic force microscopy (AFM) image of gold particles hybridised to a DNA scaffold surface arranged in regular rows in an attempt to demonstrate the possibility of DNA based nano-circuitry. Image taken from ref. 20 (B) Antibodies bound to a peptide/DNA hybrid scaffold to study peptide interactions. Image taken from ref. 20.

behind this technique is to develop a method to assemble gold particles into nanoscale circuitry for the implementation of nanoelectronics. The concept can be further developed to manipulate biological processes by incorporating binding sites within the DNA scaffold. By introducing peptides into the junctions of a DNA tiled array using a DNA-peptide conjugate, Chaput *et al.* were able to bind antibodies into a patterned arrangement in order to create program-driven peptide nano-arrays. In this way peptide-peptide interactions of the bound antibodies can be accurately studied on a nano-metre scale (Figure 6B).[22]

Whilst the use of Seeman's 'tile' approach to building DNA scaffolds from custom designed, synthetic oligonucleotides has proved to be useful in a number of studies, its use in the fabrication of larger and more complex architectures has been more challenging. Structures developed from a large number of short oligonucleotides are highly sensitive to stoichiometry and sufficient yields of

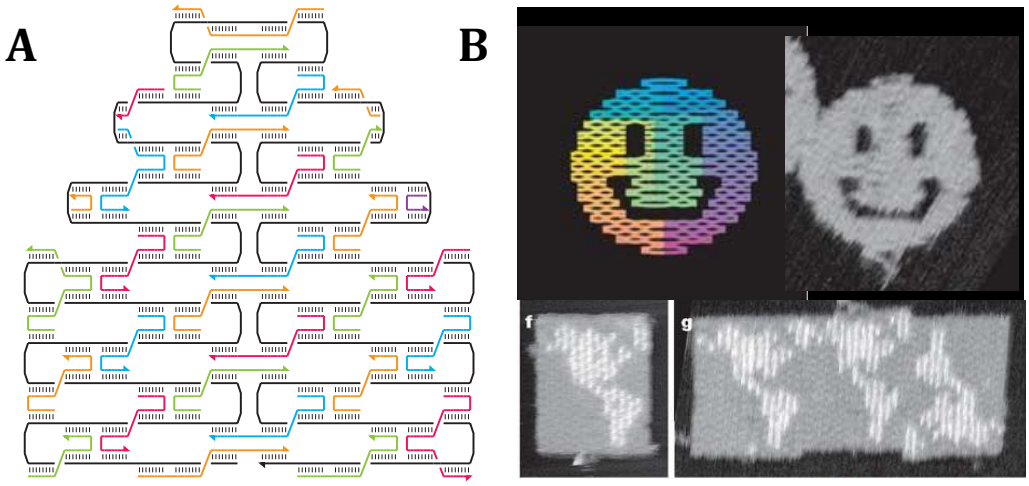


Figure 7. (A) A schematic illustration of a long sequence of viral DNA (black) folded into a pre-defined shape with the use of staple strands (various colours). Image taken from ref. 23(B) Some AFM images of folded DNA in amusing shapes, demonstrating the flexibility of the DNA origami technique. Image taken from ref. 24.

complete nanostructures are only possible through multiple reaction and purification steps. DNA origami is a technique that has sought to overcome these issues; it involves the folding of a long single strand of viral DNA using shorter "staple" strands, which bind to viral DNA in specific locations and arranging it into the desired design (Figure 7A).[23] The entire sequence of the viral DNA, along with the required design is fed into a computer, which calculates the length, sequence composition and total number of staple strands required for the correct folding to take place. Once synthesised, the staple strands are mixed with the viral DNA, then heated and cooled to initiate hybridisation and therefore folding of the nanostructure, which proceeds without any significant misfolding. Studies that have used this technique have been able to demonstrate that it can be extremely flexible, by using the same piece of viral DNA and simply changing the sequence composition of the staple strands a whole host of nanostructures can be produced, including specific two or even three dimensional shapes, (Figure 7B).[24]

The progression of DNA nano-structures into a third dimension has opened the door for further applications such as the encapsulation of reactive species. But in order to create polyhedrons of increasing complexity, scientists have had to be more creative in how they link the DNA sequences together to create the correct geometries at each of the junctions. To overcome this obstacle they have looked to the addition of synthetic groups with pre-defined geometries that can orient linear duplexes at varying degrees relative to each other. Sleiman *et al.* developed a *m*-terphenyl insertion that introduces a 60-120^{o} bend at each junction, therefore making more complex three dimensional structures possible

(Figure 8A). The Sleiman group have already demonstrated the synthesis of an array of polyhedra such as the triangular and hexagonal prisms shown in Figure 8B.[25] And due to the flexibility of DNA synthesis, the precise size of these nano-structures, and therefore their internal cavity, can be fine-tuned to fit a wide range of guests.[26]

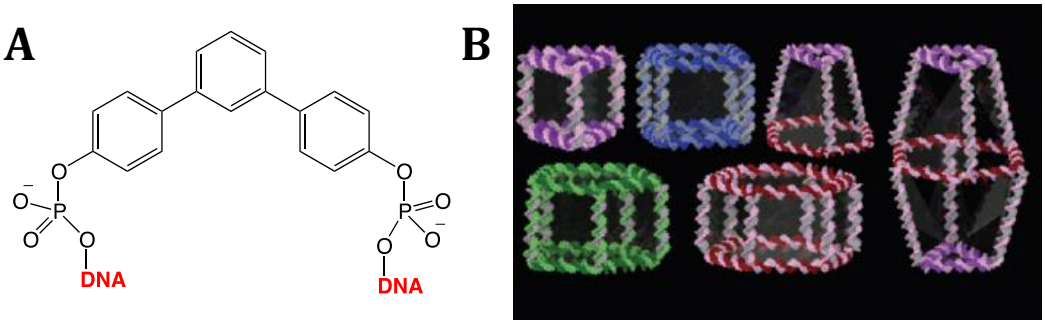
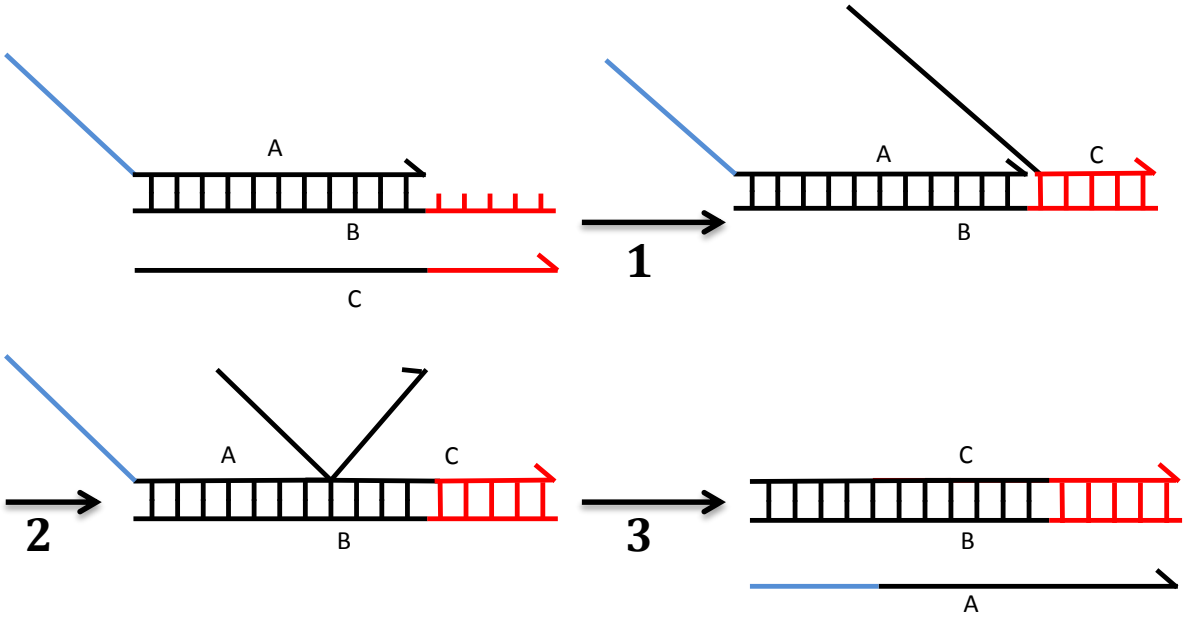


Figure 8. (A) The synthetic m-terphenyl group inserted into a DNA sequence and (B) Computer models of a variety of complex three-dimensional assemblies, the synthesis of which has been made possible using this insertion. Image taken from ref. 25

1.4 Dynamic DNA Nanotechnology

In the examples shown up to this point we have seen what is known as 'structural' DNA nanotechnology, which all have an assembly with a static, equilibrium endpoint. By making DNA nano-structures dynamic so that they can reconfigure themselves makes it possible to develop nano-devices that can be controlled by an external input.

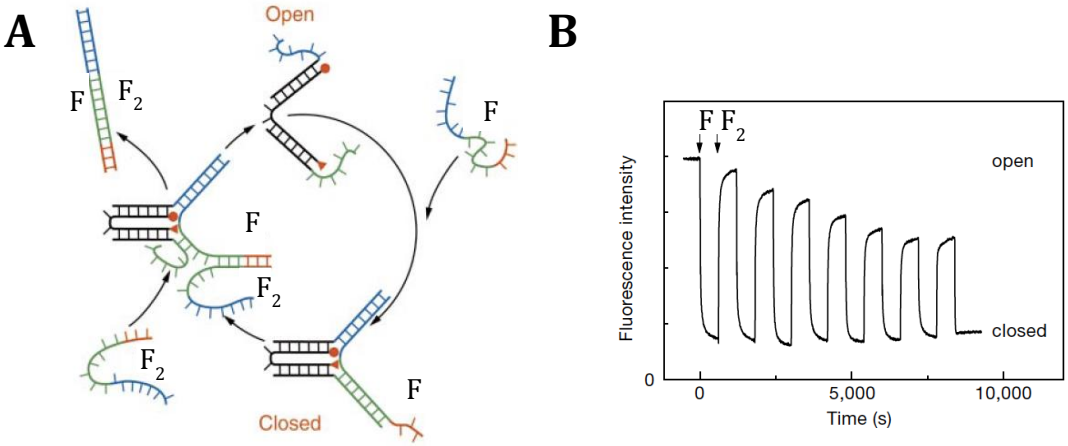
The principle method of facilitating reconfigurable structures has been sequence displacement; a process by which one single stranded segment of DNA can be swapped for another.[27] The process is initiated through complementary single stranded domains known as 'toeholds' that allow a secondary strand to lock onto the structure. Then through a process of reversible single nucleotide dissociation and hybridisation steps known as branch migration, the secondary strand displaces the original from the complex, a process that is driven by the formation of additional base pairs (Scheme 1).[28]



*Scheme 1. The process of sequence displacement, (1) The secondary sequence **C** binds to the complementary toehold domain on **B**, (2) Branch migration occurs as the base pairs begin to switch from **A** to **C**, (3) When branch migration is complete **A** is completely displaced and dissociates, leaving the stronger duplex between **B** and **C**.*

The secondary strands used in sequence displacement are often referred to as 'fuel' strands, an input that will be consumed as it drives the DNA assemblies from one state to another. The change that they induce can be utilised to carry out specific tasks and functions to create nano-scale devices. One such example

by Turberfield and Yurke known as DNA tweezers consists of two linear duplexes linked together by a flexible hinge that are capable of switching between an open and closed state (Scheme 2A). Addition of the input fuel strand (F) binds to single stranded complementary overhangs on each arm driving the tweezers to adopt the closed configuration. Following the addition of a secondary input (F2), which binds to the toehold of the fuel strand, sequence displacement is initiated, driving the tweezers back to the open state.[29]



*Scheme 2. (A) A DNA 'Tweezer' nano-device that can cycle between an open and closed state using the input fuel strand **F** and **F2**, (B) The fluorescent emission from the nano-device when functionalised with dye and quencher groups (● and ▲). Image taken from ref. 29.*

It is possible to monitor the cycle between open and closed states by the addition of dye and quencher groups in specific locations, which yield varying fluorescent emissions based on their proximity with each other (Scheme 2B).

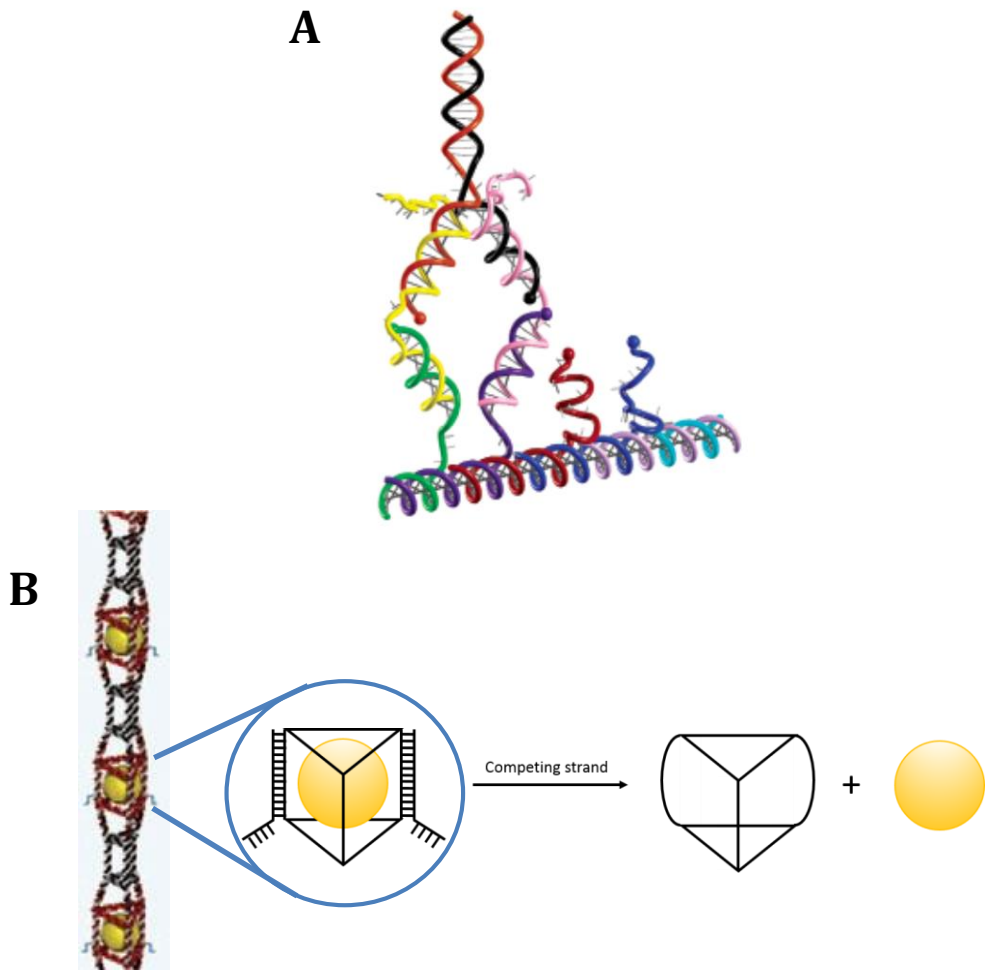


Figure 9. (A) A DNA 'walker', carrying out molecular motion by coupling to a sequence of toeholds along a DNA track through the selective addition of fuel strands. Image taken from ref. 30 (B) A dynamic 3D DNA assembly that can release encapsulated cargo through a sequence displacement reaction. In the presence of the competing strand the rigidifying sequence is removed, leaving two sides of the DNA assembly in the more flexible single stranded state allowing release of the cargo. Image taken from ref. 31.

Seeman *et al.* further demonstrated that this method of sequence displacement could be used to carry out other types of motion. His example of a nano-device that could 'walk' along a predefined trajectory was made possible by the creation of a DNA 'track'. The track included a series of toeholds that acted as steps on which the walker could bind, again through the use of expendable fuel strands (Figure 9A).[30]

The Sleiman group were also able to implement this method to create dynamic three-dimensional assemblies with variable dimensions. Their DNA nanotube design encapsulates gold nanoparticles like 'peas in a pod' using a repeating triangular prism arrangement of DNA structures (Figure 9B). Two sides of each encapsulating 'pod' utilises a short rigidifying sequence containing a short overhang domain, which can be removed in the presence of an eraser strand. On completion of the sequence displacement process two sides of the encapsulating structure are left in the non-rigid single stranded state creating a larger structure than before, thus allowing the release of the encapsulated cargo. This enhanced encapsulating property of these assemblies can facilitate the selective release of the cargo in the presence of naturally occurring sequences.[31] Therefore there is further scope to exploit this concept as a drug delivery system that will respond to pre-existing DNA/RNA sequences that are over-expressed within diseased cells as the initiator for drug release.[32] The success of this concept is dependent on being able to successfully insert the assembly into cells in such a way that they are intact, still able to function and protected from nucleases which can break down the DNA. This can be done in a number of ways including lipid based transfection, which enters the cell through endocytosis, or other techniques such as microinjection or electroporation, which bypass endocytosis. In addition to this the structure of DNA can be improved to make it less susceptible to degradation such as the PNA or LNA forms of DNA.[33,34]

1.5 Other Methods for Controlling Hybridisation.

Sequence displacement has become an invaluable tool that allows control over the configuration of certain DNA assemblies, but is it possible to create the same

effect through other forms of external inputs? It would then be conceivable that a whole host of responsive DNA based nano-devices could be created for use in nanotechnology-based solutions, in addition to the study of relevant biological processes. To explore this idea it is best to start by gaining an understanding of how hybridisation can be controlled, something that will be covered in this section along with some of the new and interesting techniques scientists have used to tackle this problem.

1.5.1 Thermal Melting of DNA

Due to the weak nature of the hydrogen bonds that interlink complementary strands of DNA, the simplest way of reversing hybridisation is by thermal denaturation. Increasing the temperature will break the intermolecular hydrogen bonds forcing the duplex to unwind and dissociate into its single stranded components. During this process the bases within each sequence will become un-stacked and in doing so an increase in UV absorption can be observed in a property known as hyperchromicity. Scientists have exploited this property to gain an understanding of nucleic acid thermodynamics by measuring the melting point (T_m) of various sequences.[35] The T_m is defined as the temperature at which 50% of the DNA strands within a sample have been dissociated. Figure 10 displays the sigmoidal curve generated during the transition from double to single stranded DNA in a thermal melting point experiment and how it can be used to determine the T_m.

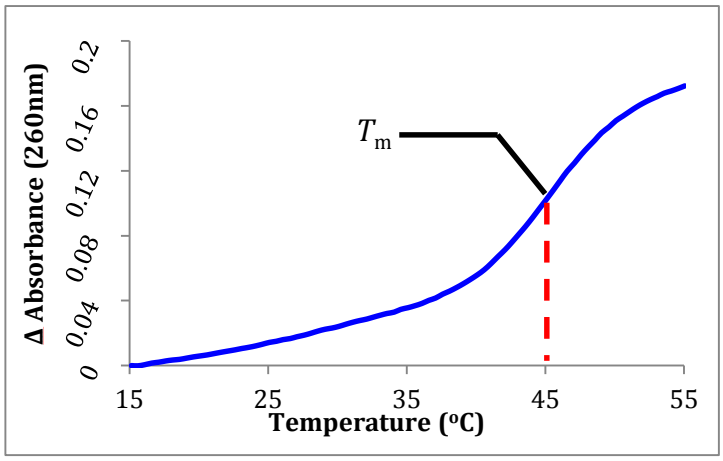


Figure 10. A variable temperature UV experiment that reveals the transition from a duplex structure to the single stranded components upon heating. The T_m is measured at the halfway point of this transition (i.e. 50% duplex) as marked on the graph (red dashed line).

The melting temperature will vary depending on the length and the specific nucleotide sequence within each sample. Put simply it is a measure of the number of hydrogen bonds within each sequence; longer sequences with a higher content of G/C base pairs (since G/C pairs have an extra hydrogen bond over A/T pairs) will create a stronger duplex and therefore possess a higher T_m.[36] The melting temperature of DNA is sensitive to the presence of base pair mismatches within a sequence, which can greatly reduce duplex stability, not only because it will have fewer hydrogen bonds than a fully matched strand, but also mismatched base pairs cause significant disruption to the duplex structure, affecting neighbouring bases. This is however not the only factor influencing duplex stability, in double stranded DNA the bases orient themselves so that they are stacked on top of each other. This is due to the aromaticity of the bases and their dipole moments, which allow for π-stacking interactions which are hydrophobic and electrostatic in nature. The degree of stabilization as a result of stacking interactions vary between sequences as some combinations of base pairs are more stable than others due to better overlap of the nucleobases. The

T_m can also be used to evaluate the inclusion of synthetic modifications and their effect upon the overall DNA structure as well as the compatibility of various duplex-binding agents.

The melting point of DNA can also alter due to other environmental factors such as the ionic strength. By increasing the concentration of counter ions that stabilise the negative charge within the sugar-phosphate backbone, the T_m can be increased to a certain extent.

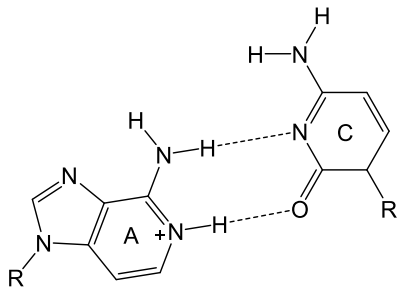


Figure 11. An A/C mismatch base pair due to protonation of adenine upon altering the pH conditions.

Furthermore the hydrogen bonds within DNA are sensitive to changes in pH and denaturing can occur in high alkaline or acidic conditions as the hydrogen bonding between the two strands is destroyed. Lowering the pH can also induce protonation of certain nucleobases, such as adenine, which can be protonated on the N(1) position with a pK_a of 3.88. The consequence of this is that the protonated adenine has the ability to hydrogen bond incorrectly to cytosine, creating A/C mismatches (Figure 11).

The correlation between T_m and base composition is approximately linear and can therefore be predicted using equation 1, which takes into account the

percentage composition of GC base pairs and a constant (X) that is dependent on salt concentration and pH (69.3 °C for 0.3 M NaCl at pH 7). More advanced algorithms such as the nearest neighbour method will also take into account the order of the bases within a sequence to give a more accurate prediction of the T_m value. This algorithm works by looking at each base pair individually, determining what bases neighbour it and then looks up the thermodynamic data from a set of predetermined values.[37] The sum of these values gives an accurate prediction of the stability of an oligonucleotide when paired with a fully matched complementary strand and can be used to confirm results or model new sequences, some versions of the algorithm will even take into account mismatches within the sequence.[38]

$$T_m = X + 0.41[\%(C + G)]\ (°C)$$

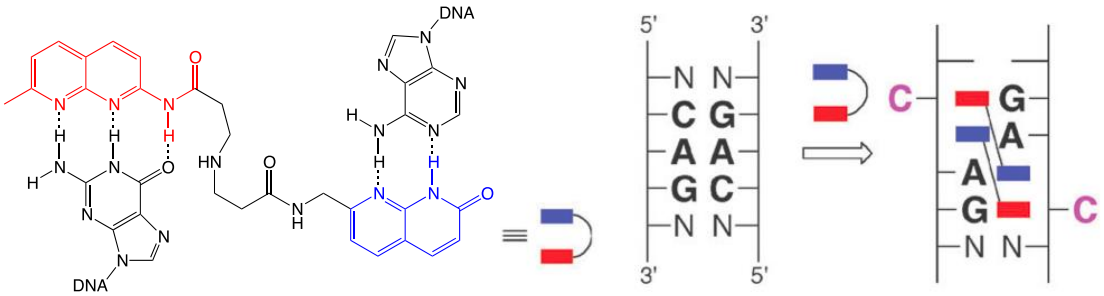
Equation 1.

Thermal denaturation could be used as an input to control the configuration of dynamic DNA assemblies by melting the structures and allowing them to reform into an alternative arrangement. However thermal denaturation can affect the entire assembly and can have drastic consequences on the functioning of the system, therefore the main challenge is to localise thermal denaturation to specific sequences within the DNA assembly. This could be done by introducing mismatches and other weak points into the desired location, allowing it to denature at lower temperatures than others, although this would have other implications. However a study by Jacobson *et al.* was able to overcome this challenge by covalently attaching oligonucleotides to metal nanocrystals that can induce localised thermal changes when exposed to an external radio frequency

field. Only those sequences within a given range of the nanocrystals will dissociate, providing control over the DNA assembly or device.[39]

1.5.2 Molecular DNA Glue

As explained previously, base pair mismatches within a DNA sequence reduce the stability of a duplex and often prevent spontaneous hybridisation from occurring. Molecular DNA glue is a concept developed by Nakatani *et al.* that aims to use this known property to their own advantage by using small ligands to bind at the mismatch site, in effect correcting the mistake and allowing the duplex to form (Scheme 3).[40]



Scheme 3. The mismatch binding ligand cross-linking two adenine and guanine nucleobases at a mismatch site to increase duplex stability. Image taken from ref. 40.

The small mismatch binding ligand (MBL) consists of two heterocyclic groups connected through a flexible alkylamino linker that are capable of forming hydrogen bonds to nucleobases. Each heterocyclic group can be designed to target a specific nucleobase and can be paired in various combinations to bind at different mismatch sites. But through experimentation it was found that a napthyridine-azaquinoline ligand bound to an AA mismatch within a CAG/CAG triad offered the greatest increase in stability. The ligand binds in a 2:1 ratio,

linking each mismatched adenine to a neighbouring guanine on the opposite strand, displacing the cytosine from that pair and effectively 'gluing' the strands together leading to a 32.8 °C increase in the T_m.[41] This has proved to be an elegant solution to gain an element of control over hybridisation, but the process is not easily reversible. To remove the MBL and revert the oligonucleotides back to the single stranded states a sufficient amount of heat must be applied to break down the ligand and the solution purified to remove the waste products.[42]

1.6 Synthetic Modification of DNA for Photo-Switched Binding

In recent research there has been a drive towards the development of techniques that make use of photo-responsive groups to regulate hybridisation. This is due to a number of distinct advantages that come with using light as an external input:

- Light is non-invasive, which is important for biological applications.
- Most photochemical transformations are rapid and lead to a change in optical properties that can be used to determine the progression from one state to another.
- There are many examples of photo-responsive groups that have reversible transformations that produce no waste products.

Photo-responsive groups within nucleic acid structures are designed to impose some sort of structural change that influences the stability of the overall duplex under isothermal conditions, shifting the melting point above and below a set temperature (ideally room temperature) in response to the absorption of light.[43] Herein is described some of those methods.

1.6.1 Caged Nucleic Acids

Caging is a term that is used to describe the incorporation of photolabile protecting groups to biologically active species; this is done at specific areas within their structure that temporarily blocks key interactions, rendering them inactive. It has been primarily used to study the time course of rapid biological processes and has already been used to study neurological processes, gene expression and cellular signalling.[44, 45]

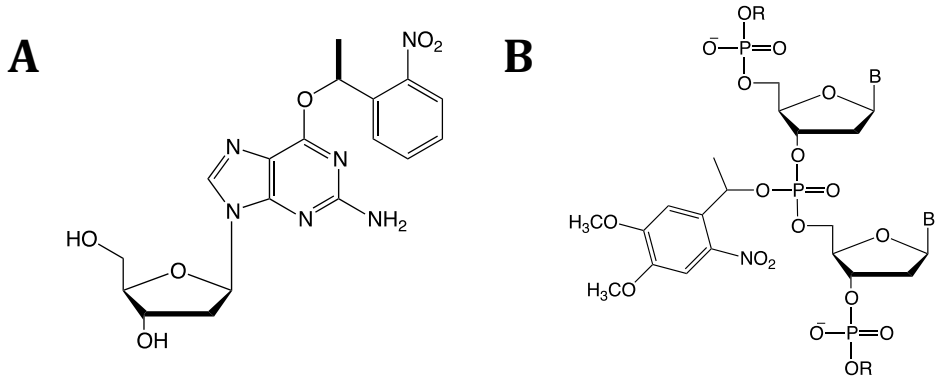


Figure 12. (A) Caging of a nucleobase using nitrobenzene (B) Caging of an oligonucleotide backbone using DMNPE

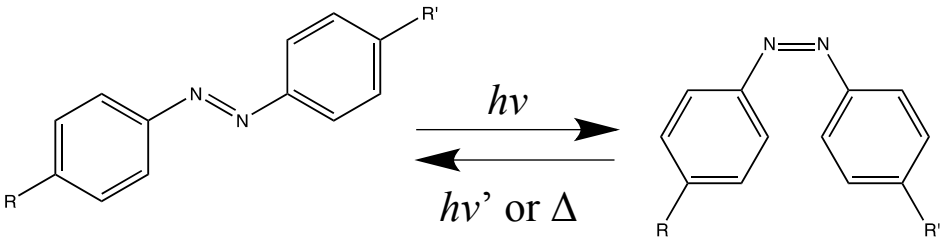
Caging has now been applied to nucleic acid structures to regulate hybridization in response to a light induced trigger and can be achieved in two ways; the protecting group can be coupled either to the heterocyclic base in order to mask the hydrogen bonding interface (Figure 12A)[46] or to the phosphate group within the backbone, neutralising the negative charge and retarding the overall structure (Figure 12B).[47] Both techniques sufficiently hinder the modified strand from forming a stable duplex until the photolabile groups are cleaved through UV irradiation.

Caging nucleic acids has provided a rapid and convenient method of regulating hybridisation with the use of light, it has already been applied in a number of

studies investigating the photo-regulation of transcription, aptamer activity and DNA regulation. But much like the molecular glue example discussed in section 1.5.2 this process is non-reversible, which can somewhat limit the possible applications.[43]

1.6.2 Photochromism

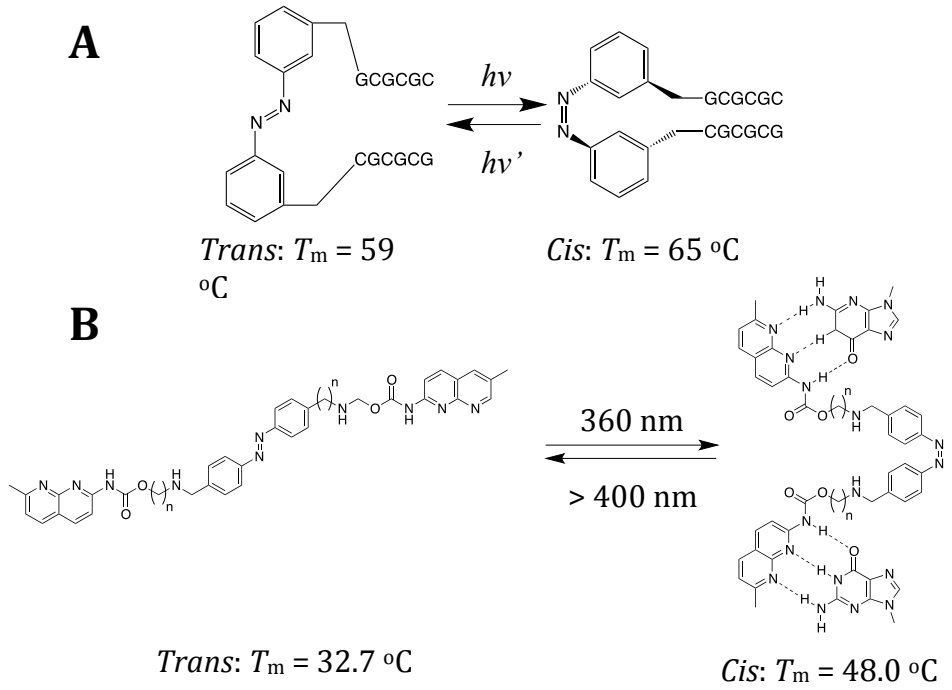
Photochromism is the reversible transformation of a chemical species upon the absorption of electromagnetic radiation. The transformation typically leads to a change in optical or structural properties that can be useful in a number of applications. Photochromic compounds that are of particular interest within this study are those that can translate a photochemical change into a structural change within the nucleic acid. A classic example is azobenzene, which undergoes *trans* to *cis* isomerisation upon the absorption of UV light between 300-400 nm (*hv*), and the process is easily reversible using a second wavelength of light > 400 nm (*hv*'), or by applying heat, (Scheme 4).



Scheme 4. The reversible, photochemical cis/trans isomerisation of a substituted azobenzene.

Functionalisation of the azobenzene at the para-position on both of the phenyl rings will enable the structural switching of the appending groups (R and R') between 180° and 60°. By coupling complementary sequences of DNA to opposite ends of the azobenzene (in this example, the meta positions), Xi *et al.*

were able to observe subtle changes in the T_m upon photo-induced isomerisation, (Scheme 5A).[48] Nakatani was also able to incorporate this process into the mismatch binding ligands in his DNA molecular glue work, in order to photo-regulate the binding to mismatch sites, (Scheme 5B).[49]



Scheme 5. (A) Alteration of duplex stability in complementary strands appended to azobenzene (B) Photo switching of a mismatch binding ligand (MBL) using azobenzene, and its binding to nucleobases on adjacent DNA strands

1.7 Incorporation of Photochromic Groups into DNA

Further research has been carried out to investigate whether photochromic compounds can be directly incorporated within an oligonucleotide, where very subtle changes in structure can potentially effect substantial changes in the binding strength. With this target in mind a number of strategies have been developed, which will be reviewed in this section.

1.7.1 Wedging of Azobenzene

Another feature of the azobenzenes discussed previously is that upon isomerisation to the *cis* form, the structure is no longer able to adopt a planar geometry due to the steric interactions between each of the phenyl groups. The work of Asanuma and Komiyama *et al.* sought to utilise this feature by focusing on the development of a non-nucleosidic linker that would allow direct insertion of an azobenzene into an oligonucleotide sequence using automated DNA synthesis.

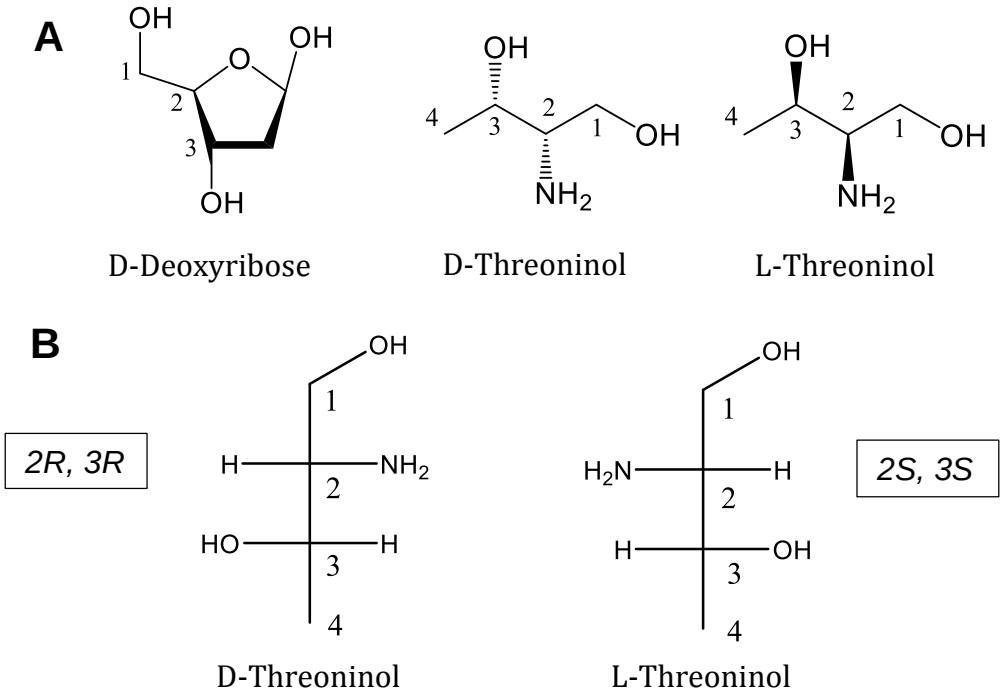


Figure 13. (A) Comparison of natural D-Deoxyribose to two enantiomers of threoninol. Note the structural similarity where primary and secondary alcohols are separated by a chain of three carbons. (B) Stereochemical assignment of D & L –threoninol using Fischer projections and the Cahn-Ingold-Prelog priority rules.

The linker chosen for this work is a diol known as D-threoninol (Figure 13), and was chosen due to a number of distinct advantages:

- Compared to natural D-ribose, D-threoninol is acyclic, but is structurally similar in that it connects a primary and secondary alcohol through a chain of three carbons.[50]
- It bears a primary amine, which will allow coupling of the azobenzene through a peptide bond.
- The configuration of the chiral centres within the D isomer of threoninol can accommodate the clockwise winding of DNA into a helical structure, unlike the L isomer.
- Threoninol can be easily prepared for DNA synthesis by the conversion to the corresponding DMT protected phosphoramidite.

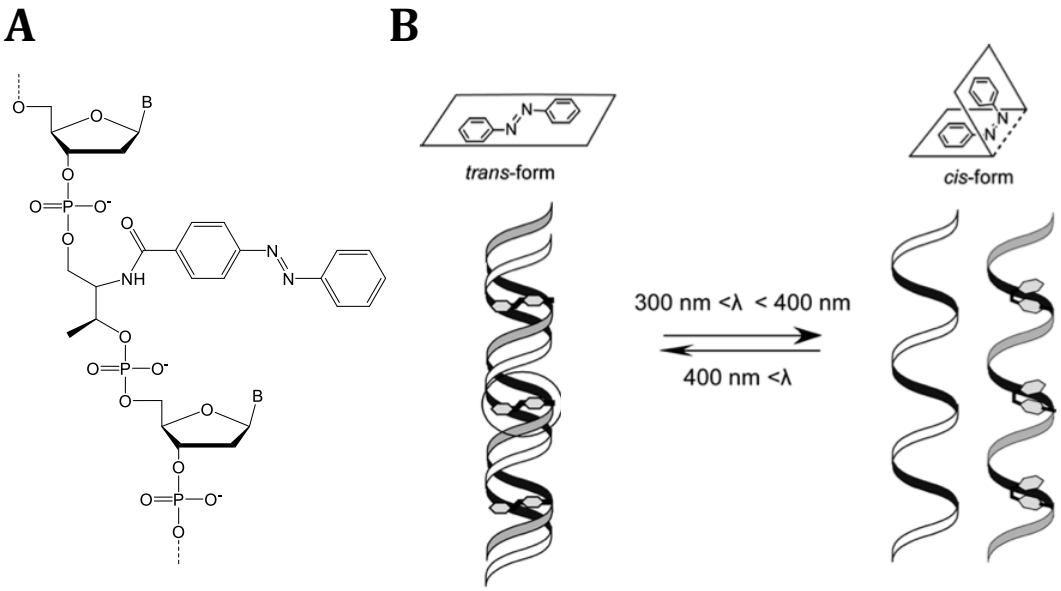
The main strategy of Asanuma's research was to insert the azobenzene modification within a range of sequences effectively making the strand an extra base pair in length. Therefore by pairing them with the shorter complementary strand the azobenzene can be 'wedged' into the cavity created by the threoninol insertion (Table 1).

Sequence name	Sequence composition
Unmodified strand	5′-CGAGTC-3′
Complementary strand	5′-TACTCG-3′
Modified strand	5′-CGA**X**GTC-3′

Table 1. A list of sequences used in Asanuma's study of photo-switched binding. **X** *marks the wedged insertion of an azobenzene group.*

It was considered that the incorporation of an acyclic structure such as threoninol into the backbone of an oligonucleotide would alter the structure, leading to destabilisation of the duplex due to an increase in flexibility. But it was found that the azobenzene was able to compensate for this unfavourable

insertion. In the planar *trans* form the aromatic phenyl groups were able to π-stack to the neighbouring bases, which rendered the unstrained duplex tight.[50]



Scheme 6. (A) The non-nucleosidic insertion of azobenzene based on a threoninol backbone structure. (B) Light induced isomerisation of azobenzene driving the transition from a duplex structure to single stranded DNA. Image taken from ref. 51.

Upon isomerisation to the *cis* form, these interactions were no longer favourable and the switch to the non-planar geometry significantly disrupts the π-stacking and hydrogen bonding of the neighbouring bases to such an extent that the duplex is destabilised (Scheme 6). In a 6-mer like that shown in Table 1, the *cis*/*trans* isomerisation was able to induce a 23.7 °C reduction in duplex stability, effectively switching between double and single stranded states using light.[51] Increasing the length of the sequences involved will require additional azobenzene modifications to achieve the same effect. It was found that the effective photo-switching of a 20-mer sequence using azobenzene would require nine modifications, once every two bases. This, in effect, means that the modified sequence is the equivalent of nine base pairs longer than the complementary strand, and although it can form a stable duplex, due to the nature of 'wedging', it

will deviate far from the B-type helical structure of natural DNA, which could make biological applications difficult.[52]

1.7.2 Stilbene Nucleotide Derivatives

One of the major disadvantages to the non-nucleosidic incorporation of a photochromic group such as azobenzene into an oligonucleotide is that whilst stacking interactions will compensate for the loss of stability, it can not take part in any hydrogen bonding interactions, nor does it sterically resemble a nucleobase. Therefore the insertion of these groups can lead to detrimental effects on the structural features of nucleic acids. The work of Maeda *et al.* attempted to use Asanuma's earlier research as a template, but instead focused on incorporating photochromic groups directly onto a nucleobase. This led to the development of stilbene-like guanosine derivatives (Figure 14),[53] which similar to azobenzenes can undergo light induced *cis*/*trans* isomerisation about a double bond attached to the 8-position of the guanine nucleobase.

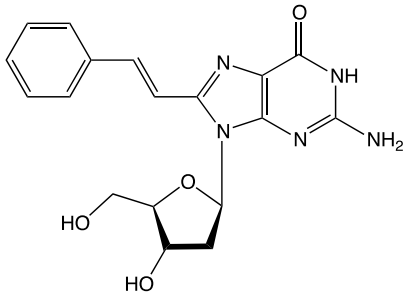


Figure 14. A Stilbene like derivative of a guanine nucleotide with photo-switched capabilities.

Insertion of these modified nucleosides into an oligonucleotide did not lead to deviation from the B-type DNA structure as observed through circular dichroism analysis, and produced melting temperatures comparable to those of unmodified

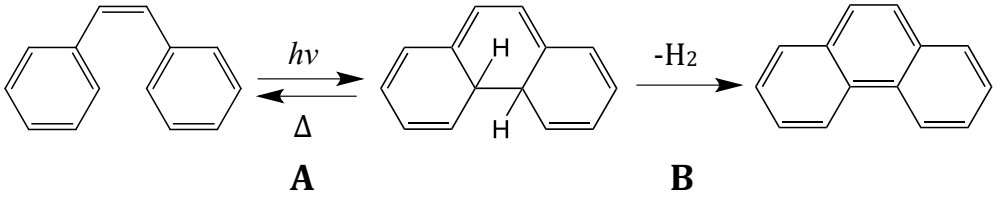
sequences. This demonstrated that the modified nucleobase was still able to take part in hydrogen bonding interactions whilst in the *trans* form.

Upon light induced isomerisation to the *cis* form, the phenyl group within the photochromic modification sterically hinders the backbone of the oligonucleotide, destabilising the duplex. It was found that with only three modifications, evenly spaced within a 20-mer, the duplex could be completely destabilised so that no T_m *trans*ition was observed. Additionally the fluorescent intensity of the photochromic nucleotides changes upon *cis*/*trans* isomerisation allowing the switch in conformational state between single stranded and duplex DNA to be monitored by fluorescence spectroscopy without any further labelling. Unfortunately this technique restricts where a modification can be placed within pre-existing sequences, as Maeda has so far only published work on modified guanosine and focused his research on the photo-regulation of G-Quadruplex sequences which are rich in guanine.[54] Ideally a range of nucleobases (A, G, C & T) would be synthesised with photochromic modifications in order to make this technique completely flexible.

1.7.3 Dithienylethenes

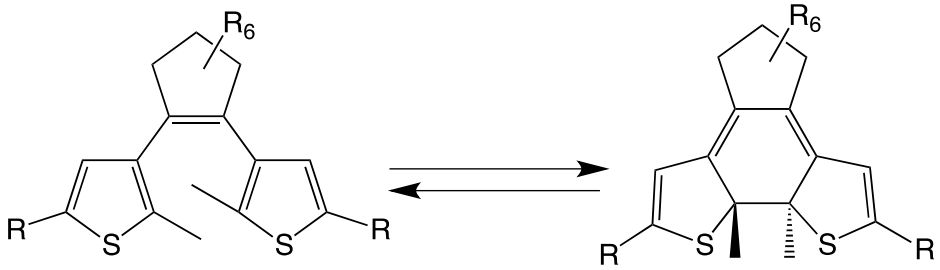
One of the main disadvantages of azobenzenes and the stilbene derivatives of Maeda described previously is that it is difficult to achieve complete conversion to the less stable *cis* isomers, with a typical yield of around 80%. This is due to what is known as a photostationary state, which occur when the two isomers of a photochromic compound have similar absorption spectra, meaning the same wavelength of UV light that drives the forward reaction can also initiate the

reverse reaction. Therefore the photostationary state defines the position of the equilibrium between each isomer, which is dependent on the relative absorbances of each isomer at a particular excitation wavelength. The optimised excitation wavelength for azobenzenes was determined to be 365 nm, whilst subsequently the complete back conversion to the *trans* isomer can be achieved under thermal conditions in the absence of light.



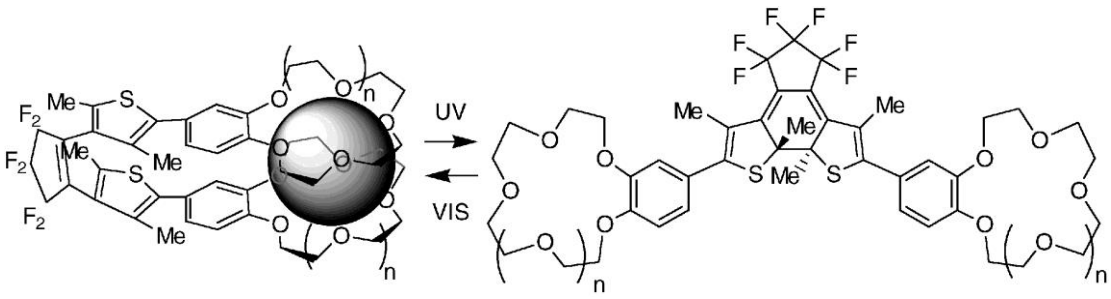
Scheme 7. (A) The photochromic, pericyclic cyclisation of stilbene to dihydrophenanthrene. (B) The irreversible oxidation of dihydrophenanthrene to give phenanthrene

Dithienylethenes are a class of compounds that have been progressively improved to have excellent photochromic properties including photo-switched products with distinct, non-overlapping absorption spectra, which eliminates the problems involving photostationary states. They were developed from stilbene, which in addition to *cis*/*trans* isomerisation is known to undergo a reversible pericyclic photo-cyclisation reaction to dihydrophenanthrene that will revert back to the stilbene form in the dark (Scheme 7A). However, in the presence of air the dihydrophenanthrene irreversibly undergoes oxidation to give phenanthrene (Scheme 7B). Substitution of the 2- and 6-positions with methyl groups suppresses this side reaction, allowing the compound to freely switch between the open and closed forms and therefore being classed as photochromic (Scheme 8). Despite the ease of photo-cyclisation of the methyl-substituted



Scheme 8. General structure of a dithienylethene molecular switch. Upon irradiation a pericyclic 6π cyclisation reaction is initiated, leading to the formation of an additional bond between the thiophene groups.

compound, it was discovered that the lifetime of the closed isomer was very short and therefore not suitable for switching applications, this was addressed by lowering the aromaticity of the system by replacing phenyl rings with heterocyclic groups such as thiophene. Finally the central double bond is locked in the cis configuration using a 5 or 6 membered ring in order to prevent *cis*/*trans* isomerisation occurring upon irradiation, rather than cyclisation.[55]

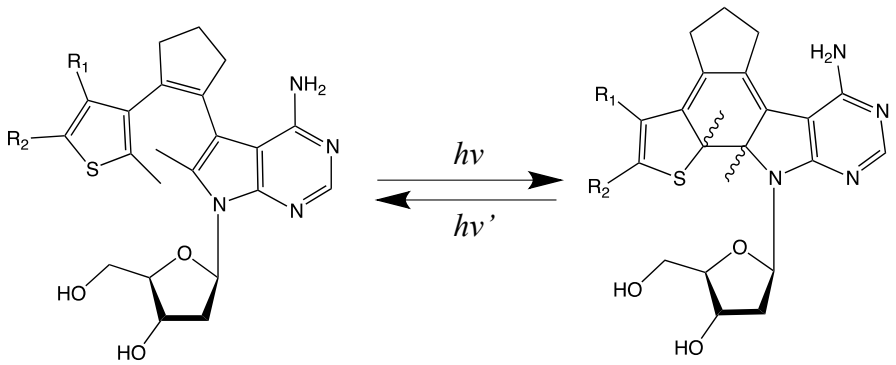


Scheme 9. Photo-responsive tweezers with a dithienylethene hinge. Switching between the open and closed forms of the dithienylethene controls the binding strength towards alkali metals, which bind to the crown ether arms. Image taken from ref. 57.

The combination of these features makes dithienylethenes highly resistant to both fatigue and thermal reversion and makes them ideal for switching purposes. They have already been utilised in a variety of applications such as photo-electrochemical switching or control over supramolecular architectures.[56,57] Work by Takeshita and Irie demonstrated a photo-responsive

tweezer design that incorporates crown ether arms onto the dithienylethene, which are capable of binding alkali metal cations. Switching between the open and closed forms of the dithienylethene regulates the proximity of each crown ether arm and therefore the overall binding strength of the tweezer structure towards metal cations (Scheme 9).[57]

Some recent work by Jäschke *et al.* has sought to utilise dithienylethenes for the photo-regulation of hybridisation using the photo-cyclisation reaction. Their research has also focused on direct functionalisation of a nucleobase and has undertaken extensive research into the modification of an adenosine nucleotide so that the ring atoms of the nucleobase are actually incorporated within this photochromic functional group (Scheme 10). It is proposed that designing the nucleotide in this way should limit the impact on structural features when incorporated within an oligonucleotide.[58]



Scheme 10. An adenine nucleotide that has been extensively modified to become incorporated within a diarylethene functional group, which is capable of performing a light induced cyclisation reaction.

In addition to these improvements, the spectral properties of diarylethenes can be tuned by changing the R_1 and R_2 substituents on the thiophene moiety. Therefore it is possible to synthesise a range of photo-switched nucleotides, each with their own unique cyclisation wavelength. It is then feasible to generate

sequences with several photo-switchable nucleobases that can be controlled independently of one another, which could potentially lead to some useful applications.[59]

Additionally the ring closure of the modified nucleobase leads to an overall extension to the delocalised system leading to a shift from a colourless starting material to a deeply coloured photo-switched product, therefore giving visual confirmation of the transformation.[58]

One of the main issues with diarylethenes is their poor solubility in aqueous solvents, therefore synthesis of a modified oligonucleotide has proved challenging. In some recent work, Jäschke has been able to post-synthetically incorporate a diarylethene modification onto an oligonucleotide using a Suzuki-Miyuara cross-coupling reaction. Results have shown that the modification did not significantly disrupt base pairing upon addition of the complementary strand, whilst the photo-cyclisation reaction remained highly efficient with good reversibility. This work has subsequently been utilised to photo-regulate the transcription process of polymerase by incorporating the modification within the T7 promoter region of a template strand and has shown promising potential for further applications.[60]

1.8 Anthracene

Anthracene is a polycyclic organic compound comprised of three fused benzene rings that is known to have interesting spectral and photochemical properties. It is a colourless compound but exhibits a blue fluorescence under UV light.

There has already been extensive research into the functionalisation of nucleic acids with anthracene to exploit its fluorescent properties for DNA/RNA sensor applications, as seen in research by Yamana and Saito.[61,62]

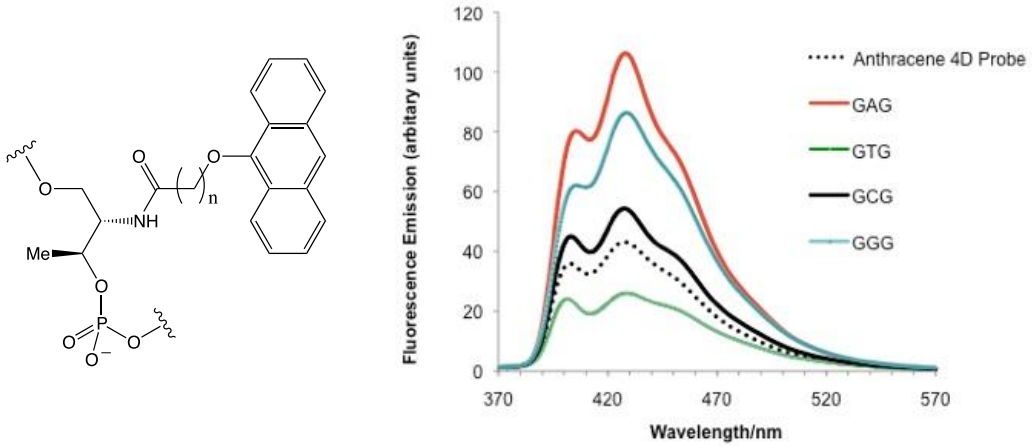
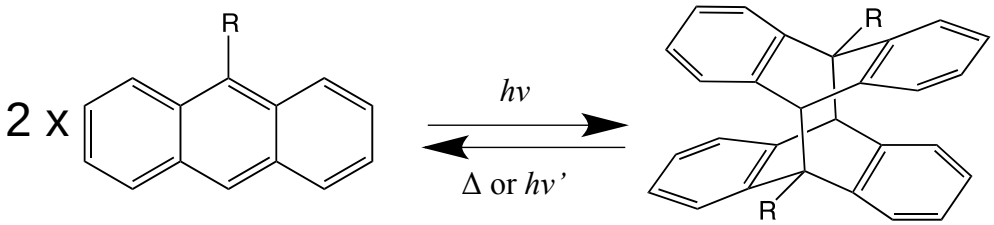


Figure 15. A non-nucleosidic anthracene insertion and the emission spectrum produced when hybridised to a range of complementary strands, each with a different base opposite the anthracene. Image taken from ref. 65.

Further research within the Tucker group led to the non-nucleosidic insertion of anthracene into nucleic acids, following on from the work of Asanuma to utilise a threoninol based backbone structure. Like *trans*-azobenzene, anthracene is a planar aromatic structure that will preferentially intercalate between neighbouring bases in the hydrophobic core of the DNA duplex through π-stacking interactions.[63] Upon pairing the modified oligonucleotide with the complementary strand it was observed that the anthracene fluorescent response could discriminate between different nucleobases placed directly opposite on the complementary sequence due to subtle differences in quenching interactions (Figure 15). The Tucker group therefore pursued several strategies to exploit this observation to detect different base compositions in an attempt to sense single nucleotide polymorphisms (SNPs) within genomic sequences.[64,65]

1.9 Anthracene Photochemistry



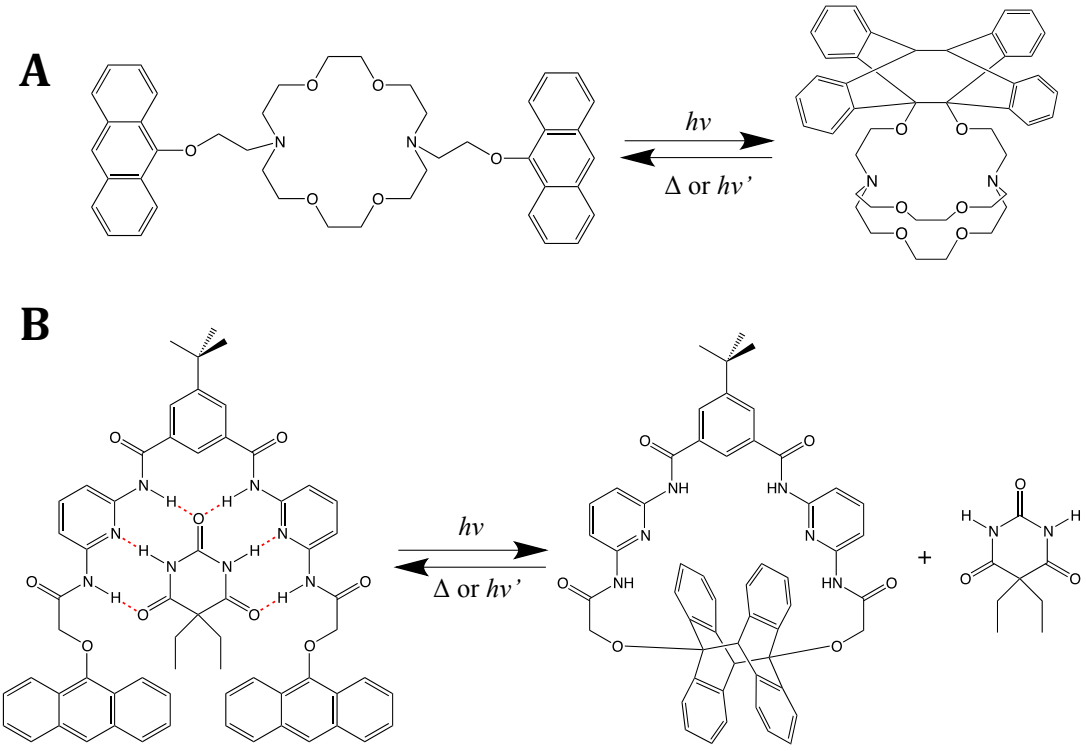
Scheme 11. The reversible, photochromic dimerisation reaction of anthracene.

Anthracene is also known to have photochromic properties, but unlike the examples shown previously the reaction is bimolecular. When anthracene in a photochemically excited state comes into contact with a secondary ground state anthracene, a 4π+4π pericyclic cycloaddition reaction occurs resulting in the formation of two new covalent bonds to create a photodimer (Scheme 11). The dimer can then easily be reverted back to the monomer forms by the application of heat or UV light of a shorter wavelength.[66]

1.10 Anthracene in Supramolecular Chemistry

Anthracene has already proved to be useful within the field of supramolecular chemistry, facilitating rapid and convenient control over supramolecular architectures. By incorporating two anthracene groups in key locations within a supramolecular receptor the dimerisation reaction can be used to alter the structure to such an extent that the binding strength towards the desired guest is significantly altered. This has already been applied to simple host-guest architectures such as the crown ether to control binding of metal ions by using the dimerisation reaction to convert it into a cryptand (Scheme 12A).[67] In contradiction to the normal macrocyclic effect that would suggest the cryptand to form stronger complexes with metal cations, it is actually the opposite in this

case, with the open crown structure giving the highest binding affinity. This is due to the bulky aromatic anthracene dimer on one of the cryptand arms, which prevents the metal cation from binding effectively. This example has also been utilised to sense different sizes of metal cations using excimer fluorescence detection. Excimers are a prerequisite for photodimerisation and arise upon the collision of an excited state and ground state anthracene, but rather than performing the pericyclic cycloaddition reaction the energy can be released in the form of fluorescence. Excimer fluorescence emission is longer than that of the excited monomer, therefore it can be detected through fluorescence



Scheme 12. (A) A crown ether modified with anthracene pendant arms for photo-controlled cryptand formation, (B) Photo-control of a Hamilton receptor using anthracene dimerisation.

spectroscopy. In the presence of larger cations such as Cs^+ the pendant anthracene arms adopt a *cis* geometry to better accommodate the ion, therefore bringing the anthracene in close proximity and making excimer formation possible. The structure adopts a *trans* geometry in the presence of smaller

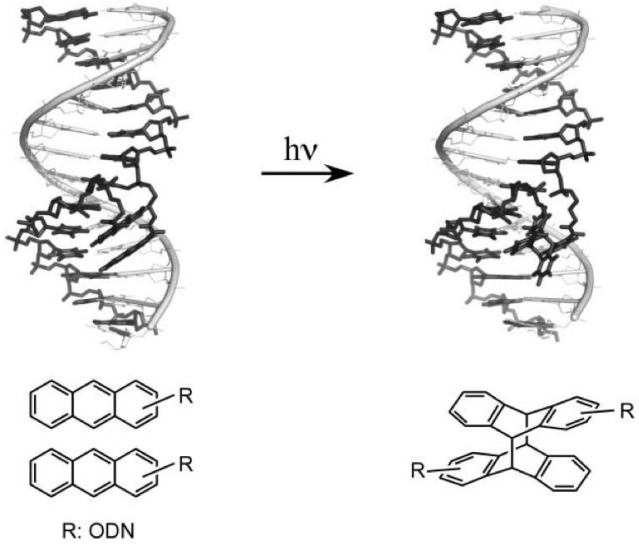
cations, with the pendant arms on opposite sides of the central macrocycle, therefore it is not possible for excimers to form.

The Tucker group has also attempted to exploit the anthracene dimerisation reaction to control more complex receptors that bind organic species such as the Hamilton receptor (Scheme 12B).[68] In this example the anthracene dimer again blocks the internal hydrogen-bonding motif, which prevents binding to the barbiturate guest, but what is interesting in this case is that they have also demonstrated the light induced ring closure of a macrocycle. The Tucker group is currently investigating whether this technique can be used to reversibly control the ring closure steps within the synthesis of interlocked structures such as rotaxanes and catenanes, which typically rely on chemical ring closure mechanisms.

1.11 Anthracene DNA conjugates

In other areas of research the focus has been the functionalisation of anthracene to separate entities in order to carry out photo-induced coupling reactions; work by Penelle *et al.* demonstrated that incorporation of anthracene into different polymers such as polystyrene would enable the formation of block co-polymers upon UV irradiation with different physical properties than the starting materials.[69] But the coupling of anthracene functionalised groups is heavily dependent on concentration as the reaction relies on the proximity of the two anthracene monomers to initiate dimerisation.

By linking anthracene to the 3' and 5' ends of two oligonucleotides and hybridising the pair to adjacent sequences within a template strand, Ihara *et al*. were able to pre-position anthracene utilising the programmable features of synthetic DNA (Scheme 13). The DNA therefore acts as a scaffold, bringing the appended anthracene groups into close proximity making it possible to achieve efficient photo-dimerisation even at low concentrations. This process is analogous to the natural process of enzymatic ligation, the joining of DNA ends through covalent bonds by enzymes such as ligase.[70]



Scheme 13. The photo-ligation of DNA using anthracene modified oligonucleotides and a DNA template. Image taken from ref. 70.

Further development of this technique has also made it possible to sense for SNPs within genomic sequences. The presence of mismatches causes disruption to the duplex structure at the ligation site and therefore the appended anthracene groups do not stack as efficiently as with fully matched. This can be observed through a reduction in the rate of photodimerisation when a mismatch is present within the template sequence. Two alleles with the location of the

polymorphism within the ligation site will display varying rates of photodimerisation, therefore allowing the SNP to be detected.[71,72]

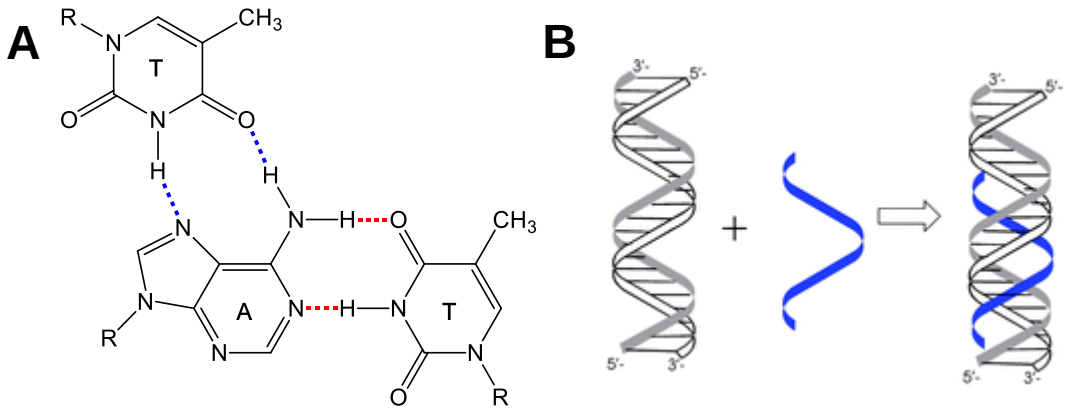


Figure 16. (A) The interaction between three oligonucleotides through Watson-Crick (red) and Hoogsteen (Blue) hydrogen bonding. (B) An oligonucleotide binding within the major groove of a B-DNA duplex through Hoogsteen hydrogen bonding, forming the triplex structure.

Ihara and co-workers have also demonstrated that the technique can be applied to other types of DNA secondary structure, one such example is the DNA triple helix. Triple helixes form due to additional hydrogen bonding donor and acceptors on purine (Pu) nucleobases, which allow a secondary oligonucleotide to bind (Figure 16). Pyrimidines (Py) do not possess this feature and therefore a pre-requisite for triplex formation is a homopurine region within one strand of the duplex DNA and a homopyrimadine region within the complementary sequence. The complementary sequence binds within the major groove of the duplex and will orient itself either parallel or antiparallel to the purine strand of the duplex depending on the sequence composition (Figure 17A). Antiparallel sequences contain the Pu.PuPy motif, whereas parallel sequences contain the Py.PuPy motif. Parallel triplexes are more stable at lower pH's (below pH 6) due to the protonation of cytosine, which allows for additional Hoogsteen hydrogen bonding (Figure 17B).

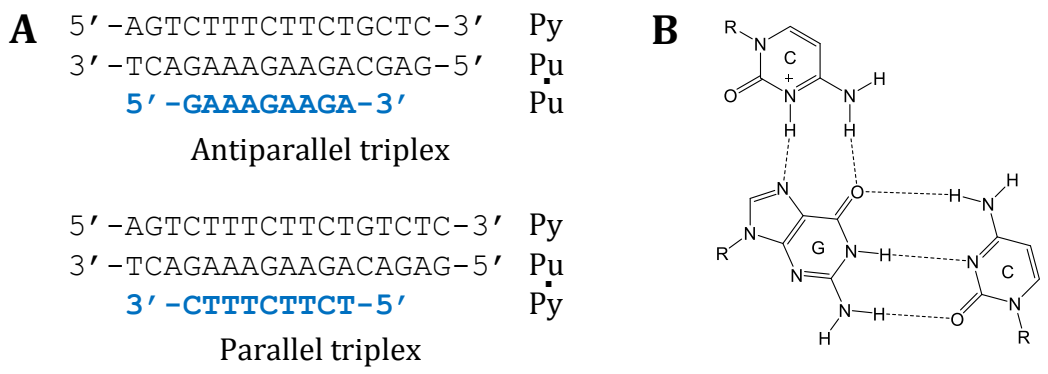


Figure 17. (A) Examples of triplex forming sequences including the antiparallel (Pu.PuPy motif) and parallel (Py.PuPy motif) forms. (B) Protonation of N(3) of cytosine, which allows stable Hoogsteen bonding in the parallel triplexes at low pH.

Ihara's research utilises a bimolecular triplex, which consists of an oligonucleotide that has been doubly modified with anthracene on both the 3' and 5' ends. The sequence of this oligonucleotide contains two mirror repeat heptamer segments connected by a four nucleotide loop, which has been specifically designed to fold into the triplex structure in the presence of a target strand (Figure 18). The formation of the triplex structure brings the appended anthracene groups into close proximity, allowing the photochemical reaction to be initiated, the result of which connects the two ends of the oligonucleotide, creating a circular piece of DNA.[73] Circular DNA is important in biological

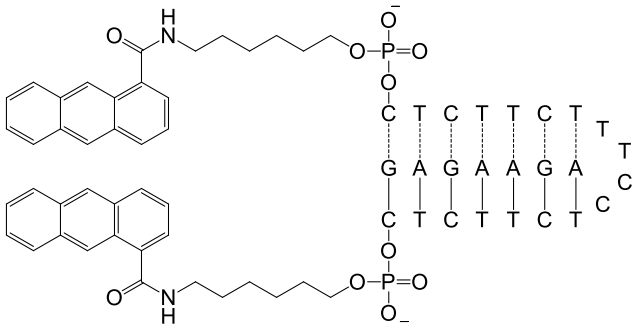
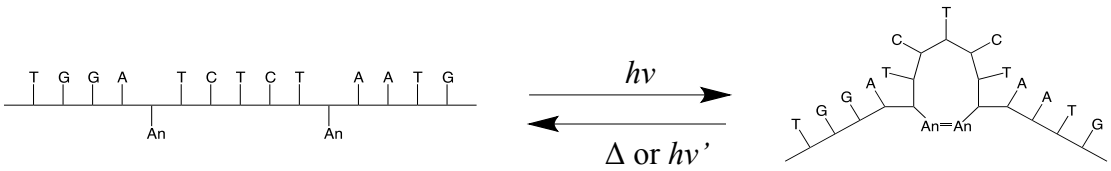


Figure 18. An anthracene modified bimolecular triplex developed by Ihara for use in generating circular DNA.

applications and can be found in bacteria and viruses, therefore this technique could prove useful as a reversible method of generating circular DNA for use in various studies.

1.12 Project concept

The main aim of this research is to design a variety of double-tagged anthracene oligonucleotides that will allow the photochromic properties of anthracene to be applied to the control of duplex structures. It is hypothesised that the photodimerisation reaction will impose a structural change on the DNA double helix that will dramatically alter the overall stability of the resulting duplex when paired with the complementary sequence (Scheme 14). The anthracene will be incorporated into DNA using the same non-nucleosidic linker based on threoninol that was utilised by previous members of the group as well as in research by Asanuma and Komiyama with their work on azobenzenes.



Scheme 14. A double anthracene-tagged oligonucleotide and a schematic representation of the structural change that is induced upon photodimerisation, (An = anthracene tag).

Due to the convenience of automated, solid phase DNA synthesis it will be easy to control the position of each anthracene relative to each other by simply changing their corresponding positions within an oligonucleotide sequence. By varying the number of bases in between each anthracene tag, and therefore the distance between them it is proposed that the extent of the structural change

that will be imposed on DNA through photodimerisation can be controlled/customised.

The project will initially focus on the photo-control of standard double helical structure of B-DNA, but will also experiment with more complex forms of DNA structure, such as the stem loop or G-quadruplex. With careful consideration of each of the hybridised/folded structures, the anthracene will be placed at key locations within the sequence that will hopefully cause maximum structural change upon dimerisation.

The efficiency of each sequence to control hybridisation through photodimerisation will be analysed through a combination of techniques such as UV and variable temperature UV spectroscopy, gel electrophoresis and HPLC. The structural features will also be analysed by circular dichroism spectroscopy, to give an indication of the extent of change that is induced upon photodimerisation.

1.13 References

1. Lehn, J.-M., *Nobel Prize Interview*, **2001**, http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1987/lehn-interview.html (accessed on 8/12/13)
2. Cragg, P., *Supramolecular Chemistry*, Springer, Netherlands, **2010**, Chapter 1, 1-48.
3. Pedersen, J., *J. Am. Chem. Soc.*, **1967**, *89* (26), 7017-7036.
4. Lehn, J., *Angew. Chem. Int. Ed.*, **1988**, *27* (1), 89-112.
5. Hof, F.; Rebek, J., *Proc. Natl. Acad. Sci. USA*, **2002**, *99* (8), 4775-4777.
6. Fisher, E., *Berichte der Deutschen Chemischen Gesellschaft,* **1894**, *27*, 2985-93.
7. Chang, S.; Hamilton, A., *J. Am. Chem. Soc.*, **1988**, *110* (4), 1318-1819.
8. Hamilton, A.; Pant, N.; Muehldorf, A., *Pure & Appl. Chem.*, **1988**, *60* (4), 533-538.
9. Bissell, R. A.; Calle, E.; de Silva, A. P.; de Silva, S. A.; Gunaratne, H. Q. N.; Habib-Jiwan, J.-L.; Peiris, S. L. A.; Rupasinghe, R. A. D. D.; Samarasinghe, T. K. S. D.; Sandanayake, K. R. A. S.; Soumillion, J.-P., *J. Chem. Soc. Perkin Trans. 2*, **1992**, 1559-1564.
10. Lehn, J., *Proc. Natl. Acad. Sci. USA*, **2002**, *99* (8), 4763-4768.
11. Blackburn, M. Gait, M. Loakes, D. and Williams, D., *Nucleic Acids in Chemistry and Biology*. RSC Publishing, Cambridge, **2006**, Chapter 2, 13-76.
12. Watson, J.; Crick, F., *Nature*, **1953**, *171*, 737-738.
13. Kool, E. T.; *Chem. Rev.*, **1997**, *97*, 1473-1487.
14. Richmond, T.; Davey, C., *Nature*, **2003**, *423*, 145-150.
15. Stulz, E.; Clever, G.; Shionoya, M.; Mao, C, *Chem. Soc. Rev.*, **2011**, *40*, 5633-5635.
16. Seeman, N. C., *J. Theor. Biol.*, **1982**, *99*, 237-247.
17. Seeman, N. C., *Nature*, **2003**, *421*, 427-431.
18. Chen, J.; Seeman, N. C., *Nature*, **1991**, *350*, 631-633.
19. Pinheiro, A.; Han, D.; Shih, W.; Yan, H., *Nature Nano.*, **2011**, *6*, 763-772.
20. Pinto, Y.; Le, J.; Seeman, N.; Musier-Forsyth, K.; Taton, T.; Kiehl, R., *Nano Lett.*, **2005** *5* (12), 2399-2402.
21. Yan, H., Park, S.; Finkelstein, G.; Reif, J.; LaBean, T., *Science*, **2003**, *301*, 1882-1884.
22. Williams, B.; Lund, K.; Liu, Y.; Yan, H.; Chaput, J., *Angew. Chem. Int. Ed.*, **2007**, 46, 3051-3054.
23. Rothemund, P. W. K., *Nature*, **2006**, *440*, 297-302.
24. Woo, S.; Rothemund, P. W. K., *Nature Chem.*, **2011**, *3*, 620-627.
25. Aldaye, F.; Palmer, A.; Sleiman, H., *Science*, **2008**, *321*, 1795-1799.
26. McLaughlin, C.; Hamblin, G.; Sleiman, H., *Chem. Soc. Rev.*, **2011**, *40*, 5647-5656.

27. Zhang, D.; Seelig, G., *Nature Chem.*, **2011**, *3*, 103-113.
28. Meselson, M. S.; Radding, C. M., *Proc. Nat. Acad. Sci. USA*, **1975**, *72* (1), 358-361.
29. Yurke, B.; Turberfield, A. J.; Mills, A. P.; Simmel, F. C.; Neumann, J. L., *Nature*, **2000**, *406*, 605-608.
30. Sherman, W. B.; Seeman, N. C., *Nano Lett.*, **2004**, *4*, 1203-1207.
31. Lo, P. K.; Karam, P.; Aldaye, F. A.; McLaughlin, C. K.; Hamblin, G. D.; Cosa, G.; Sleiman, H. F., *Nat. Chem.*, **2010**, *2*, 319-328.
32. Lo, P. K.; Metera, K. L.; Sleiman, H. F., *Curr. Opin. Chem. Biol.*, **2010**, *14*, 597-607.
33. Wittung, P.; Nielson, P. E.; Buchardt, O.; Egholm, M.; Nordén, B., *Nature*, **1994**, *368*, 561-563.
34. Singh, S. K.; Nielsen, P.; Koshkin, A. A.; Wengel, J., *Chem. Commun.*, **1998**, 455-456.
35. SantaLucia, J.; Hicks, D., *Annu. Rev. Biophys. Biomol. Struct.*, **2004**, *33*, 415-440.
36. Breslauer, K. J.; Frank, R.; Blöcker, H.; Marky, L., *Proc. Natl. Acad. Sci. USA*, **1986**, *83* (11), 3746-3750.
37. Breslauer, K. J.; Frank, R.; Blöcker, H.; Marky, L., *Proc. Natl. Acad. Sci. USA*, **1986**, *83*, 3746-3750.
38. For an example of T_m prediction software and associated algorithms see: biophysics.idtdna.com (accessed on 8/12/13)
39. Hamad-Schifferil, K.; Schwartz, J.; Santos, A.; Zhang, S.; Jacobson, J., *Nature*, **2002**, *415*, 152-155.
40. Nakatani, K.; Sando, S.; Saito, S., **2001**, *19* (1), 51-55.
41. Nakatani, K.; Hagihara, S.; Goto, Y.; Kobori, A.;Hagihara, M.; Hayashi, G.; Kyo, M.; Nomura, M.; Mishima, M.; Kojima, C., *Nat. Chem. Biol.*, **2005**, *1*, 39-43.
42. Dohno, C.; Nakatani, K., *Chem. Soc. Rev.*, **2011**, *40*, 5718-5729.
43. Brieke, C.; Rohrbach, F.; Gottschalk, A.; Mayer, G.; Heckel, A., *Angew. Chem. Int. Ed.*, **2012**, *51*, 8446-8476.
44. Mayer, G.; Heckel, A., *Angew. Chem. Int. Ed.*, **2006**, *45*, 4900-4921.
45. Tang, X.; Dmochowski, I. J., *Mol. BioSyst.*, **2007**, *3*, 100-110.
46. Mayer, G.; Kröck, L.; Mikat, V.; Engeser, M.; Heckel, A., *ChemBioChem*, **2005**, *6* (11), 1966-1970.
47. Ghosn, B.; Haselton, F. R.; Gee, K. R.; Monroe, W. T., *Photochem. Photobiol.*, **2005**, *81*, 953-959.
48. Wang, Q.; Gao, S.; Zhou, K.; Chen, W.; Niu, C.; Xi, Z., *Chin. J. Chem.*, **2009**, *27* (8), 1582-1588.
49. Dohno, C.; Uno, S. N.; Nakatani, K., *J. Am. Chem. Soc.*, **2007**, 129 (39), 11898-11899.
50. Kashida, H.; Liang, X.; Asanuma, H., *Curr. Org. Chem.*, **2009**, *13*, 1065-1084.

51. Asanuma, H.; Takarada, T.; Yoshida, T.; Tamaru, D.; Liang, X. G.; Komiyama, M., *J. Am. Chem. Soc.*, **2001**, *40*, 2671-2673.
52. Asanuma, H.; Matsunaga, D.; Komiyama, M., *Nucleic Acids Symp. Ser.*, **2005**, *49*, 35-36.
53. Ogasawara, S.; Maeda, M., *Angew. Chem. Int. Ed.*, **2008**, *47*, 8839-8842.
54. Ogasawara, S.; Maeda, M., *Angew. Chem. Int. Ed.*, **2009**, *48*, 6671-6674.
55. Irie, M., *Chem. Rev.*, **2000**, *100*, 1685-1716.
56. Gilat, S. L.; Kawai, S. H.; Lehn, J.-M., *J. Chem. Soc., Chem. Commun.*, **1993**, 1439-1442.
57. Takeshita, M.; Irie, M., *J. Org. Chem.*, **1998**, *63*, 6643-6649.
58. Singer, M.; Jäschke, A., *J. Am. Chem. Soc.*, **2010**, *132*, 8372-8377.
59. Jäschke, A., *FEBS Lett.*, **2012**, *586*, 2106-2111.
60. Cahová, H.; Jäschke, A., *Angew. Chem. Int. Ed.*, **2013**, *52*, 3186-3190.
61. Yamana, K.; Aota, R.; Nakano, H., *Tet. Lett.*, **1995**, *36* (46), 8427-8430.
62. Saito, Y.; Motegi, K.; Bag, S. S.; Saito, I., *Bioorg. Med. Chem.*, **2008**, *16* (1), 107-113.
63. Moran, N.; Bassani, D. M.; Desvergne, J. –P.; Keiper, S.; Lowden, P. A. S.; Vyle, J. S.; Tucker, J. H. R.; *Chem. Commun.*, **2006**, 5003-5005.
64. Duprey, J. –L. H. A.; Bassani, D. M.; Hyde, E. I.; Ludwig, C.; Rodger, A.; Vyle, J. S.; Wilkie, J.; Zhao, Z.; Tucker, J. H. R., *Supramol. Chem.*, **2011**, *23*, 273.
65. Duprey, J. –L. H. A.; Zhao, Z.; Bassani, D. M.; Manchester, J.; Vyle, J. S.; Tucker, J. H. R., *Chem. Commun.*, **2011**, *47*, 6629-6631.
66. Bouas-Laurent, H.; Castellan, A.; Desvergne, J.; Lapouyade, R., *Chem. Soc. Rev.*, **2001**, *30*, 248-263.
67. McSkimming, G.; Tucker, J. H. R.; Bouas-Laurent, H.; Desvergne, J., -P.; Coles, S. J.; Hursthouse, M. B.; Light, M. E., *Chem. Eur. J.*, **2002**, *8* (15), 3331-3342.
68. Molard, Y.; Bassani, D. M.; Desvergne, J. P.; Moran, N.; Tucker, J. H. R.; *J. Org. Chem.*, **2006**, *71* (22), 8523-8531.
69. Goldbach, J. T.; Russell, T. P.; Penelle, J., *Macromolecules*, **2002**, *35*, 4271-4276.
70. Ihara, T.; Fujii, T.; Mukae, M.; Kitamura, Y.; Jyo, A., *J. Am. Chem. Soc.*, **2004**, *126*, 8880-8881.
71. Mukae, M.; Ihara, T.; Tabara, M.; Jyo, A., *Org. Biomol. Chem.*, **2009**, *7*, 1349-1354.
72. Arslan, P.; Ihara, T.; Mukae, M.; Jyo, A., *Anal. Sci.*, **2008**, *24*, 173-176.
73. Arslan, P.; Jyo, A.; Ihara, T., *Org. Biomol. Chem.*, **2010**, *8*, 4843-4848.

Chapter 2: Synthesis

2.1 Introduction

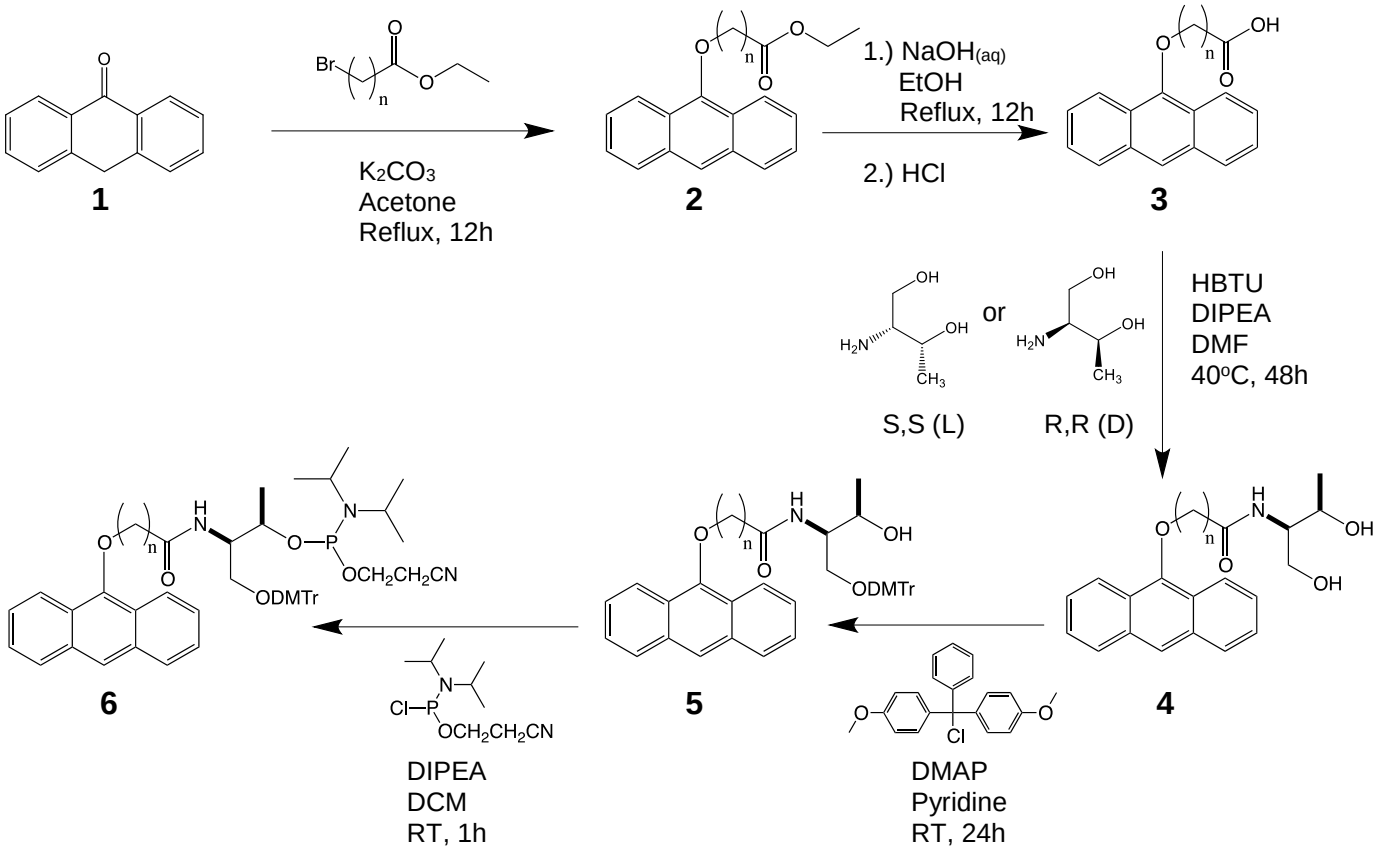
This section describes the methods used for the synthesis of an anthracene monomer, including the techniques in which different linkers are produced. It then discusses automated solid phase DNA synthesis and how this is used to incorporate an anthracene monomer into an oligonucleotide sequence, as well as the methods to purify and characterise the finished oligonucleotides.

The development of the synthesis steps described herein was carried out by previous members of the group; N. Moran and J.-L. Duprey, which was used for their work involving fluorescent DNA sensors. The work in this project largely follows the same procedures, but has been adapted in order to create a new type of anthracene monomers as well as to incorporate certain improvements.

The focus of this project is to develop oligonucleotide sequences containing multiple anthracene modifications for the purpose of anthracene photodimerisation. Therefore this chapter will discuss how this can be achieved and the implications involved during the synthesis and characterisation of the desired oligonucleotide.

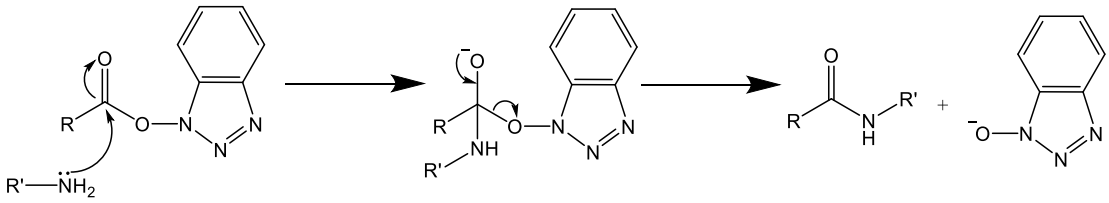
2.2 Synthesis Route of Anthracene Phosphoramidites

The synthetic strategy for the development of an anthracene phosphoramidite monomer is displayed in Scheme 1.[1] It starts from with the commercially available anthrone **1**, which was initially coupled with a brominated ester through a base-induced nucleophilic substitution reaction to give the anthracene ester **2**. A range of brominated esters were used in this step, differing from each other by the number of carbons between the bromine and carbonyl groups, denoted by the letter *n*. Anthracene can undergo oxidation to anthraquinone, which is not emissive and photochemically inactive. To minimise this process, this and all further reactions were carried out under an argon atmosphere and in the absence of light.



Scheme 1. Synthesis route for the production of n = 4, 6 and 7 anthracene monomers starting from the commercially available anthrone.

The ester **2** was then converted to the carboxylic acid **3** through a saponification reaction using aqueous sodium hydroxide in ethanol followed by precipitation using hydrochloric acid. In the next step the carboxylic acid **3** was coupled to either D or L threoninol via a peptide bond using the coupling reagent HBTU. The coupling reagent is used to activate the carboxylic acid by creating an activated ester using a benzotriazole, which provides an excellent leaving group. This favours the nucleophilic attack, lowering the activation energy and allowing the reaction to progress more rapidly (Scheme 2.).[2,3]



Scheme 2. Formation of a peptide bond using an active ester with a benzotriazole leaving group.

Threoninol is a vital component for the incorporation into the DNA framework, primarily because it is a diol, which is required to create the two phosphodiester linkages during DNA synthesis. Another important feature of threoninol is the presence of two stereogenic centres that give rise to the isomers D-threoninol (2R,3R) and L-threoninol (2S,3S), which are enantiomers of each other and both commercially available in isomerically pure forms. Therefore, the diol product is synthesised as the isomerically pure compounds **4(D)** and **4(L)**.[4,5]

Standard automated DNA synthesis requires the 5'-hydroxy group to be protected to prevent successive coupling reactions, therefore the primary alcohol on the threoninol is protected with an acid labile dimethoxytrityl (DMTr) group to give **5(D/L)**.[6] It was found in earlier research by Duprey that the DMTr group is

selective for the primary alcohol, this is due to steric clashes between the bulky DMTr and the methyl group on the secondary alcohol, which was confirmed by a number of 2D NMR techniques.[7]

In the final step **5(D/L)** undergoes a phosphitylation on the secondary alcohol using a chlorophosphitylating reagent, to create a phosphite triester **6(D/L)**. This is a key component in oligonucleotide synthesis that will create the initial phosphodiester linkage to the 5' alcohol on the preceding nucleotide. The phosphitylation reagent is commercially available with the necessary cyanoethyl protecting group and diisopropylamine leaving group needed for the oligonucleotide coupling reaction.

Anthracene monomers where n = 4 and 6 were synthesised in agreement with previous work carried out by both Moran and Duprey. However the requirement in this research for longer linkers so that photodimerisation can occur across larger distances led to the synthesis of a new n = 7 linker. This was done by simply using commercially available ethyl 8-bromooctanoate in the initial step to give the additional carbon that was required. After successive reactions the final phosphoramidite product was fully characterised by high-res mass spectrometry, as well as ^{31}P and ^{1}H NMR spectroscopy, which indicated a successful synthesis of the desired product (Figure 1).

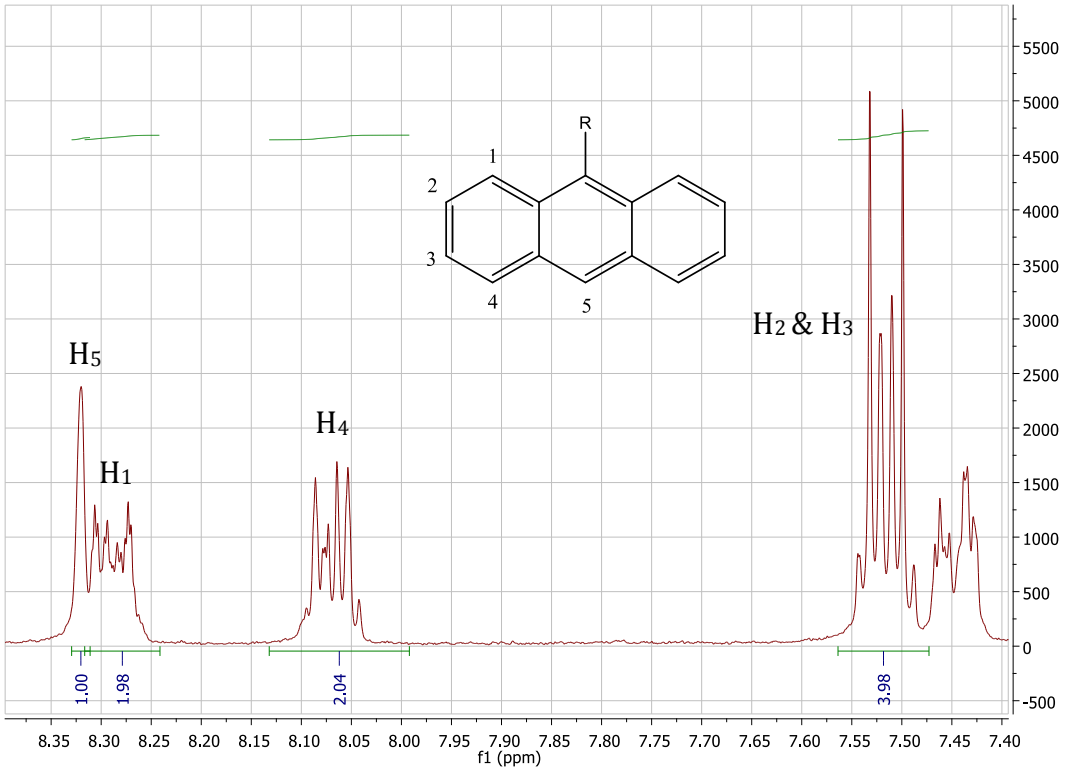


Figure 1. Characteristic NMR signals of anthracene substituted in the 9-position taken from the spectrum of an n = 7 anthracene phosphoramidite.

2.3 Design of Preliminary Sequences

The sequence chosen to adapt with the anthracene modification has been used extensively within the group in earlier work; in this way results obtained in this project can easily be compared with those previously reported. The sequence is 15 base pairs in length, which after being paired with its complementary sequence would have an almost equal ratio of A:T / G:C base pairs (8/7) that will provide a stable duplex structure in the presence of one another.[8]

Due to the ease of oligonucleotide synthesis, each sequence can be incorporated with any number of anthracene tags and in any desired position. To test the photodimerisation reaction an initial sequence (**S1**) was created, with two anthracene tags located centrally within the sequence separated by a single

nucleotide, as shown in Figure 2, where the location of each anthracene tag is denoted with the use of an **X**. Each anthracene tag is incorporated into the sequence using the **4D** linker, i.e. a four-carbon alkyl spacer combined with the D isomer of threoninol.

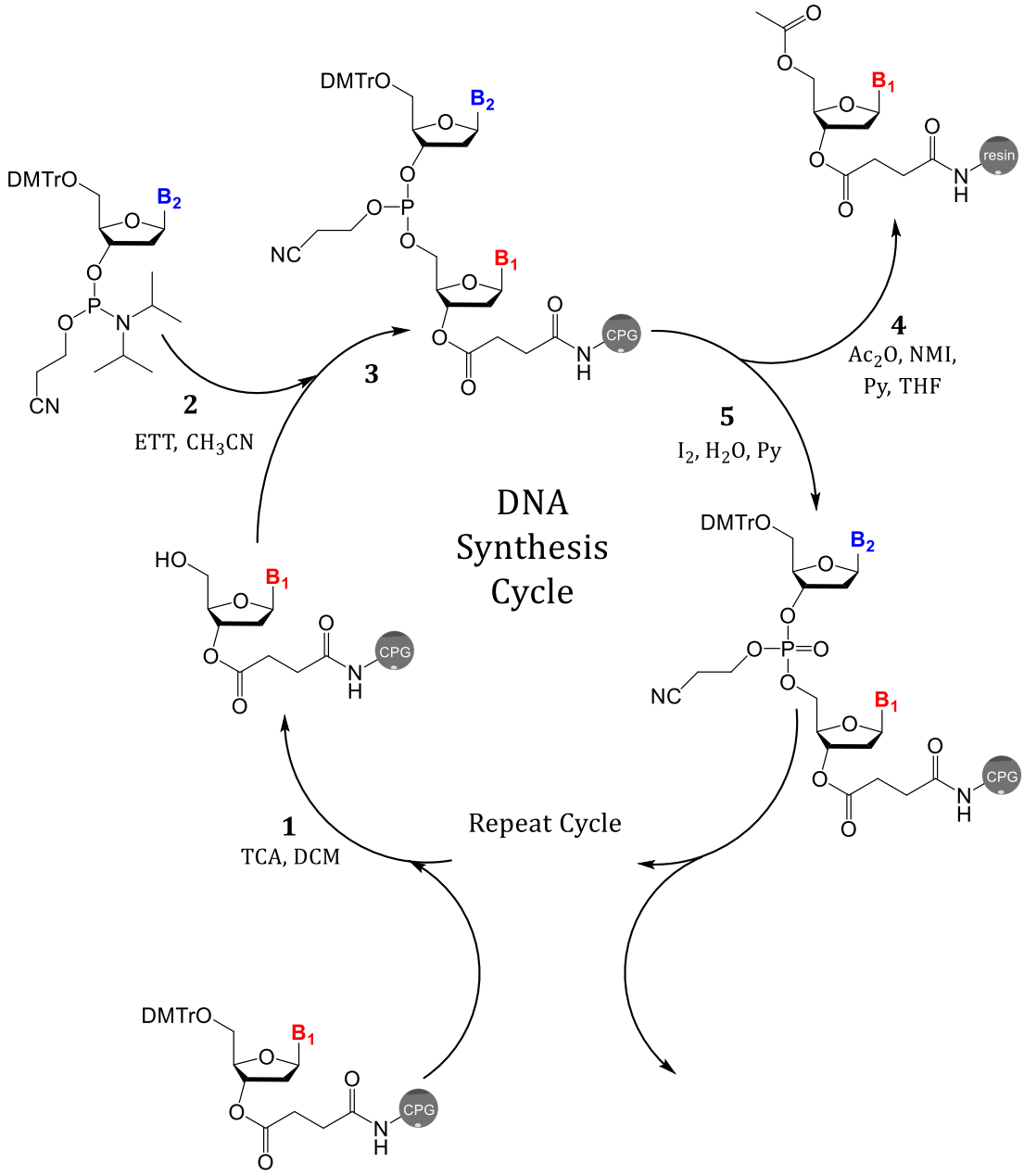
Unmodified Sequence:	5′-TGGACTCTCTCAATG-3′
Anthracene tagged sequence (**S1**):	5′-TGGACT**X**T**X**TCAATG-3′

*Figure 2. Sequence composition of the oligonucleotide and the corresponding anthracene modified sequence (**S1**)*

2.4 DNA Synthesis

The initial development of an efficient, automated process for the synthesis of oligonucleotides was carried out by Caruthers and co-workers in the early 1980s, which began with their work on phosphite triester chemistry.[9] It involves the sequential addition of mononucleotides to a deoxynucleoside that has been covalently attached to a solid support, which can provide high coupling yields and minimal side reactions. By placing the initial, support bound deoxynucleoside into a column and fitting to an automated DNA synthesiser, reagents can be added stepwise in a repeating cycle, extending the oligonucleotide sequence without the loss of product.[6]

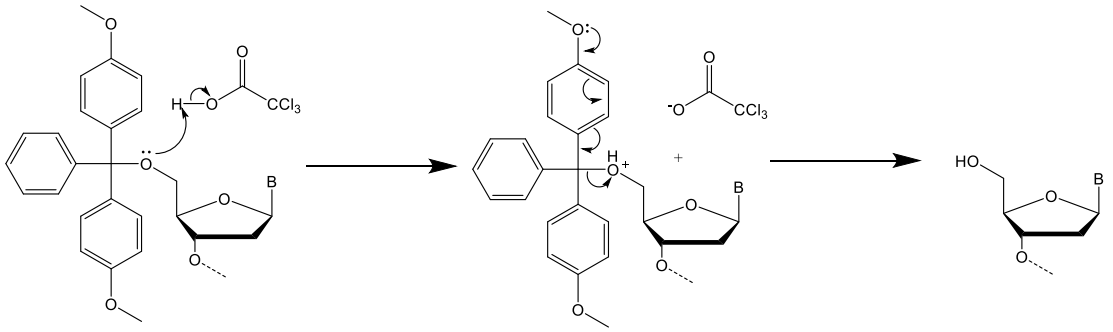
Synthesis of oligonucleotides was carried out on a 1 μM scale using an Applied Biosystems 394 DNA/RNA synthesiser, which is capable of introducing the anthracene modification at the desired position within an oligonucleotide sequence based on the synthesis cycle shown in Scheme 3.



Scheme 3. Automated, solid-phase DNA synthesis cycle.

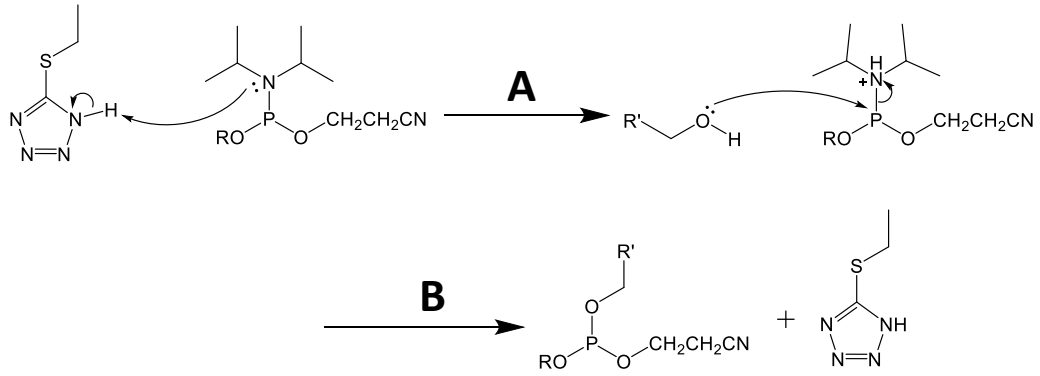
1. **Detritylation**: The initial base of the sequence, which is covalently attached to a controlled pore glass (CPG) solid support through a succinyl linker, is treated with trichloroacetic acid to remove the 5'-DMTr group, giving a free hydroxyl (Scheme 4). The positively charged DMTr cation released in this step can be measured quantitatively using a conductivity

cell, allowing the coupling efficiency to be monitored throughout the synthesis.



Scheme 4. Detritylation of the 5'-hydroxyl group using trichloroacetic acid.

2. **Activation**: Phosphoramidites of each of the required nucleotides are dissolved in acetonitrile and are mixed with a tetrazole activator (Scheme 5A) and then delivered simultaneously to the synthesis column.[10]



*Scheme 5. (**A**) Activation of phosphoramidite using 5-Ethylthio-1H-tetrazole, and (**B**) coupling with the preceding 5'-hydroxyl.*

The exocyclic amino groups on the nucleobases are sensitive to the conditions used during oligonucleotide synthesis; therefore the phosphoramidites of adenine, guanine and cytosine are available with these groups already protected, shown in Figure 3. As thymine has no exocyclic amino group, the phosphoramidite does not require protection.[11]

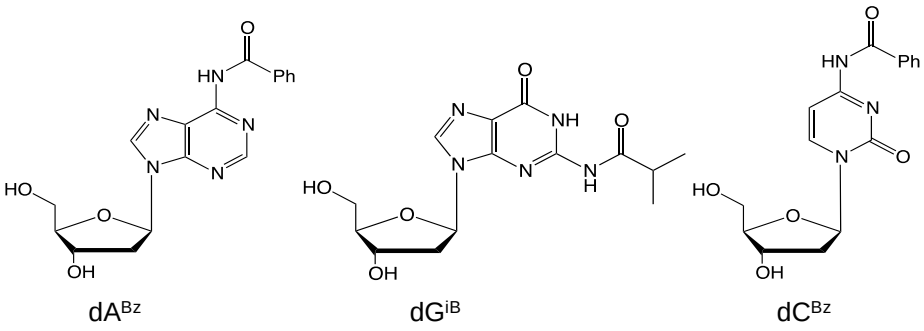
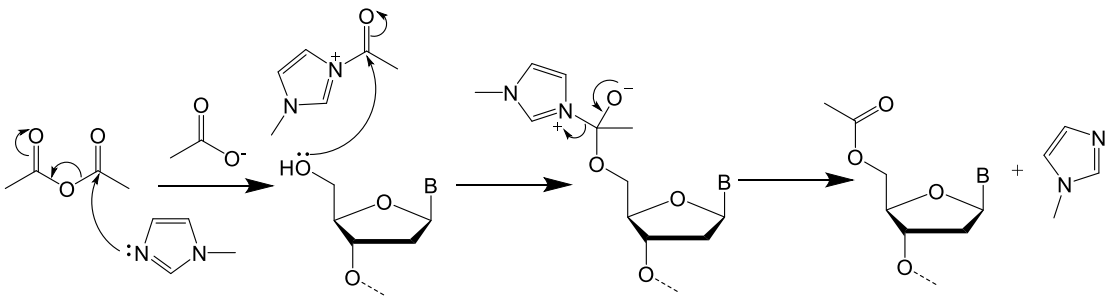


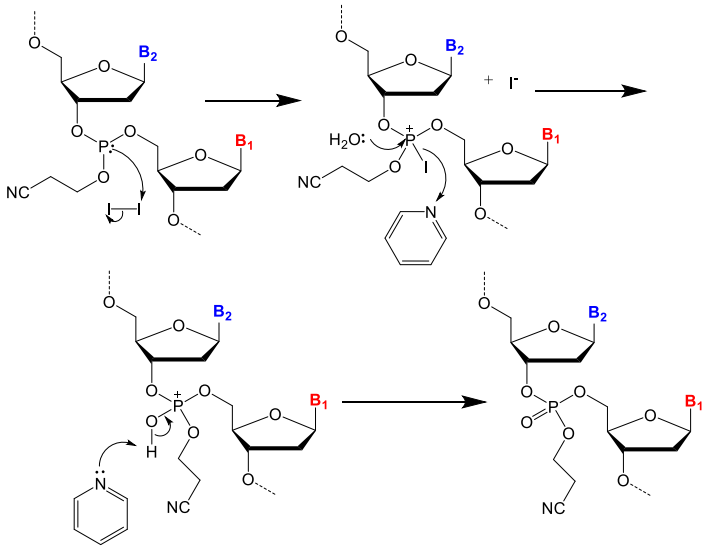
Figure 3. Nucleotides of adenine, guanine and cytosine where the exocyclic amino groups have been protected.

3. **Coupling**: Once the activated phosphoramidite is delivered to the synthesis column it couples with the primary alcohol of the preceding nucleotide to extend the oligonucleotide sequence (Scheme 5B). After the coupling is complete unreacted materials are removed by thorough washing of the solid support with acetonitrile.
4. **Capping**: Although the coupling is rapid and efficient, there may be a small percentage of 5'-OH that remain unreacted and it is important to cap these in order to prevent the generation of sequences with base deletions. This is done by performing an acetylation using a capping mixture containing acetic anhydride (Scheme 6). The capping mixture also contains pyridine which keeps the conditions basic enough that the acetic acid produced in this step does not initiate premature detritylation.



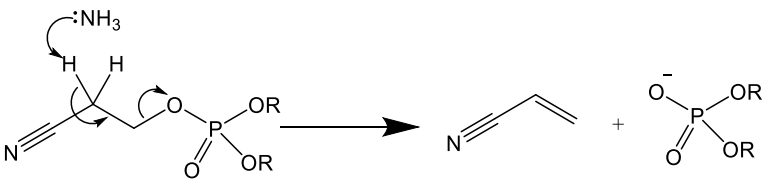
Scheme 6. Acetylation of unreacted 5'-hydroxyl groups using acetic anhydride used in the capping step of solid phase oligonucleotide synthesis.

5. **Oxidation**: The penultimate step is the oxidation of the phosphite triester to the more stable phosphate triester, which is performed by adding a mixture of iodine, pyridine and water in THF (Scheme 7). The synthesiser will then feed back into another reaction cycle until the whole oligonucleotide sequence is complete.



Scheme 7: Oxidation of a phosphite triester to the more stable phosphate triester.

6. **Cleavage and Deprotection**: The completed oligonucleotide is cleaved from the CPG support using concentrated ammonia and the collected solution is heated to 55 ^{o}C overnight to remove the protecting groups on both the nucleobases and the phosphates, leaving the finished product (Scheme 8). Ammonia is then removed by evaporation to give a crude product which is then redissolved in DNA grade water, ready to undergo HPLC purification.



Scheme 8: Post synthetic deprotection of the cyanoethyl protecting group.

2.5 DNA Purification

The pure oligonucleotide is isolated through reverse phase preparative HPLC using a 0.1 M TEAA (pH 7) buffer and an increasing gradient of acetonitrile over a period of 40 minutes ('DMT On' method, as described in the experimental section). Oligonucleotides are normally synthesised and removed from the synthesiser with the final 5'-DMTr in place, which increases the hydrophobicity and leads to a longer retention time, allowing the full-length sequence to be isolated from the deletion oligonucleotides that have been capped. In this case, the anthracene modification behaves in the same way, therefore the 5'-DMTr group can be removed on the synthesiser and the sequence can still be isolated with relative ease. A UV-Vis detector was used to monitor the output from the HPLC column, which is set to the absorption maxima of DNA (260 nm), but monitoring a second wavelength at the characteristic anthracene absorption band (360 nm) aided in the identification of the product peak.

After the main fraction was collected, the elution solvent was removed *in vacuo* using a rotary evaporator, re-dissolved in DNA grade water and passed through a NAP column, which is a form of size exclusion chromatography that can remove salts and other low molecular weight impurities. To assess the purity, all sequences are run on analytical RP-HPLC using the same gradient conditions as the preparative method, (Figure 4).

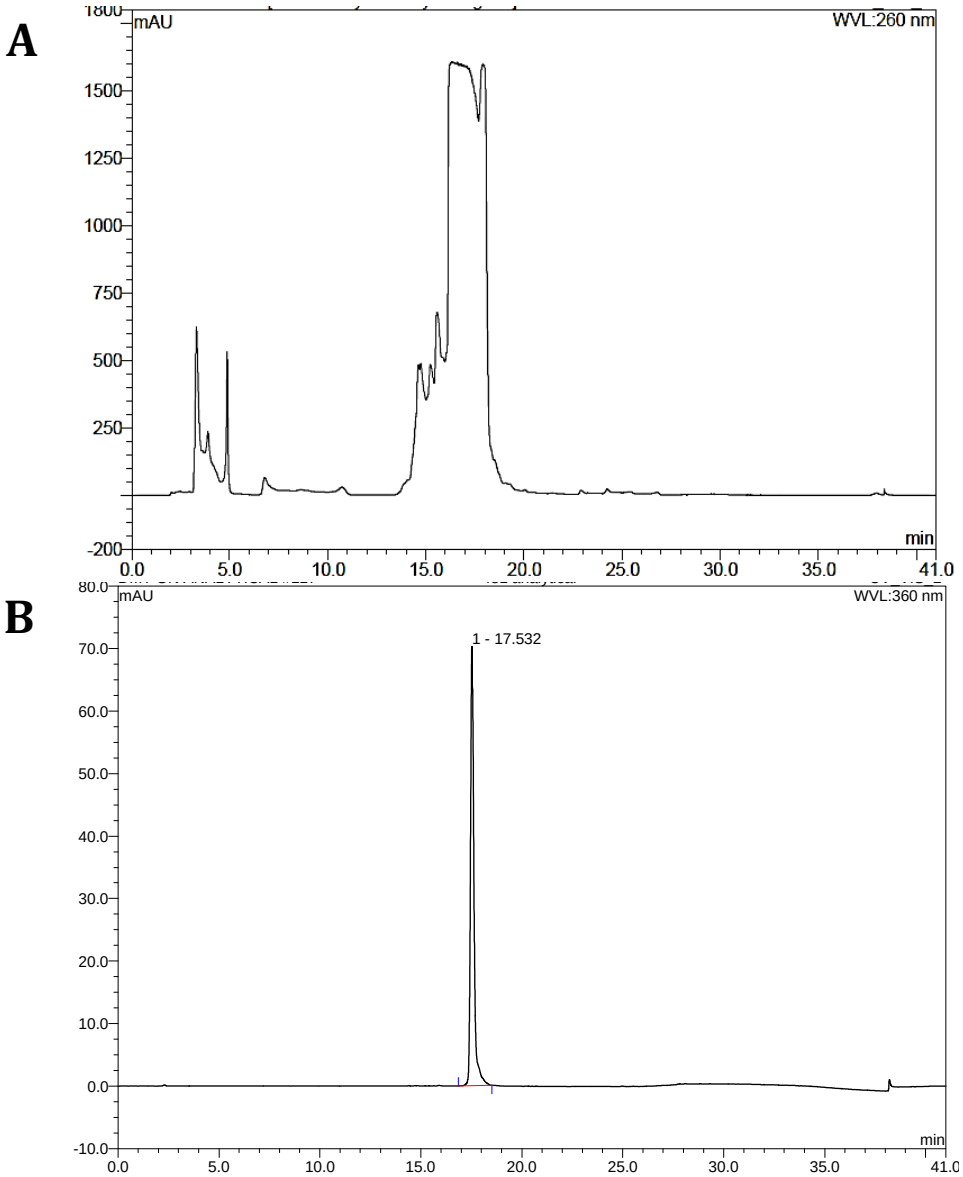


*Figure 4. (A) A preparative HPLC chromatogram of the crude **S1** oligonucleotide following the completed DNA synthesis. (B) Analytical chromatogram of pure **S1** after HPLC purification.*

2.6 Oligonucleotide Characterisation

The principle method of characterising an oligonucleotide is electrospray mass spectrometry, where the resulting peak is compared with a predicted value determined from the structure of the sequence.[13] Due to the fact that oligonucleotides have multiple charges on the phosphate backbone, the raw data

will containing multiple peaks depending on the charge of each species and requires conversion to a molecular mass profile. This is done using maximum entropy processing, an algorithm built into mass spectrometry software that can be used to mathematically produce the parent mass spectrum. The resulting spectrum provides a single peak representing the full mass of the oligonucleotide, which can be more easily compared to the predicted values of each sequence. The obtained mass spectrum of sequence **S1** matched that of the predicted value (4545.0). There were some additional peaks observed with a slightly lower mass than that of the desired product, but they were determined to be nucleobase fragments that are produced during the process of obtaining the mass spectrum (4411 – adenine, 4435 – cytosine).

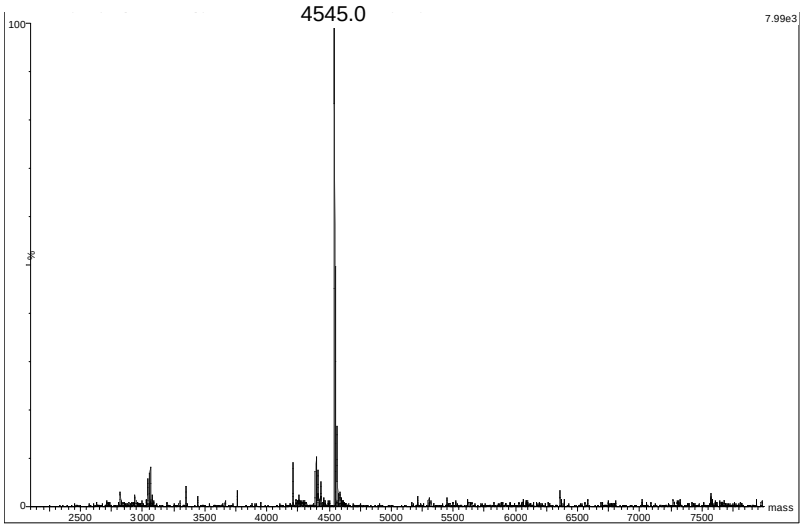


*Figure 5. ES⁻ Mass spectrum displaying the full mass of the desired product (**S1**) after maximum entropy processing. Predicted mass: 4851.4, found: 4545.0*
*Sequence **S1**:* TGGACT**X**T**X**TCAATG

2.7 Quantification

Oligonucleotides are easily quantified through UV spectroscopy due to the absorption band between 200 and 280 nm. By measuring the maximum absorbance at 260 nm the concentration of oligonucleotide can be calculated using the Beer Lambert Law, (Figure 6).

$$A = \varepsilon c l$$

A = Absorbance
ε = molar extinction coefficient
c = molar concentration
l = path length

Figure 6. The Beer Lambert law and its definitions.

The molar extinction coefficient of each oligonucleotide can be obtained using online calculators, which sum the coefficient values for each of the individual nucleotides, and use a procedure known as the 'nearest neighbour' method to take into account the order of the nucleotides and provide the final value of ε.[14,15]

However, anthracene has a strong absorption within the region of 260 nm, which must therefore be taken into account when calculating the molar extinction coefficient. Duprey was able to calculate the contribution of anthracene to the

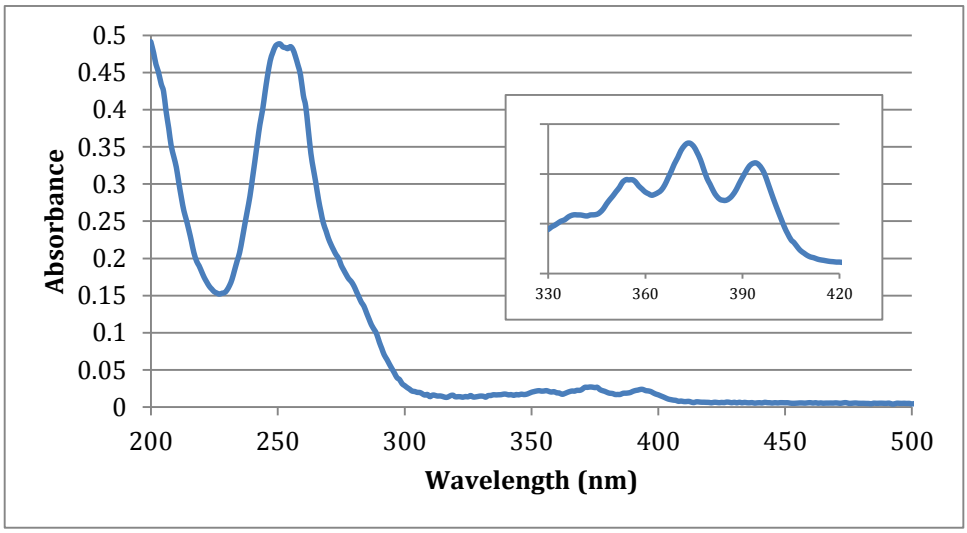


Figure 7. Anthracene modified oligonucleotide UV absorption spectrum. Featuring a DNA band at 260 nm and an anthracene band at 360 nm (enlarged inset).

value of ε by measuring the absorbance of known quantities of an anthracene monomer and applying the Beer Lambert law to produce a coefficient of an individual anthracene (51,444 M^{-1} cm^{-1}), which can be added to the value produced by the online calculator.[7]

Figure 7 shows a typical spectrum of a purified oligonucleotide, which displays the characteristic band of the oligonucleotide (maxima 260 nm) overlapping with the low energy (300-400 nm) anthracene band (^1L_a). A second fine structure high energy anthracene band (^1B_b) is seen between 320 and 420 nm (inset). The ^1L_a and ^1B_b nomenclature originates from the perimeter free electron orbital (PFEO) theory as described by Birks, used to model the electronic transitions in aromatic compounds.[16] Figure 8 displays how this theory can be applied to anthracene, depicting the weak dipole oscillations of the L_a and L_b states along with the stronger dipole oscillations of the B_a and B_b states. These electronic transitions are what make up the bands seen in the UV spectrum of anthracene.

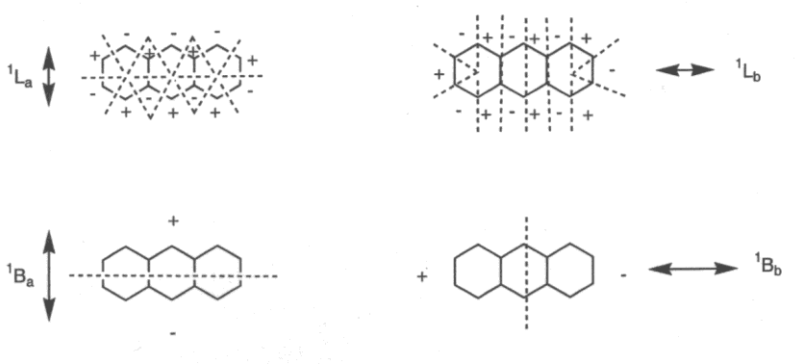


Figure 8. PFEO theory applied to anthracene.

2.8 Storage

It was discovered that under ambient lighting conditions it was possible to observe the formation of photodimers in certain sequences over time. This was

seen by the appearance of new peaks in the analytical HPLC, therefore all oligonucleotides were re-purified, placed into light-protected containers and split into aliquots to minimise the exposure to ambient light. All sequences are then stored in the freezer at -20^{o}C in DNA grade water.

2.9 References

1. Duprey, J.-L. H. A.; Zhao, Z.; Bassani, D. M.; Manchester, J.; Vyle, J. S.; Tucker, J. H. R., *Chem. Commun.*, **2011**, *47*, 6629-6631.
2. Han, S.-Y.; Kim, Y.-A.; *Tetrahedron*, **2004**, *60*, 2447-2467.
3. Knorr, R.; Trzciak, A.; Bannwarth, W.; Gillessen, D., *Tetrahedon Lett*, **1989**, *30*, 1927-1930.
4. Kashida, H.; Liang, X.; Asanuma, H., *Curr. Org. Chem.*, **2009**, *13*, 1065-1084.
5. Asanuma, H.; Toda, T.; Murayama, K.; Liang, X.; Kashida, H., *J. Am. Chem. Soc.*, **2010**, *132*, 14702-14703.
6. Gait, M. J., *Oligonucleotide Synthesis, a Practical Approach*, Oxford: IRL Press, **1984**.
7. Duprey, J.-L. H. A., (**2010**) *Studies on Anthracene Tagged Oligonucleotides*, PhD Thesis, University of Birmingham, UK.
8. Moran, N.; Bassani, D. M.; Desvergne, J.-P.; Keiper, S.; Lowden, P. A. S.; Vyle, J. S.; Tucker, J. H.R., *Chem. Commun.*, **2006**, 5003-5005.
9. Matteucci, M. D.; Caruthers, M. H., *J. Am. Chem. Soc.*, **1981**, *103*, 3185-3191.
10. Berner, S.; Mühlegger, K.; Seliger, H., *Nucleic Acids Res.*, **1989**, *17*, 853-864
11. Froehler, B. C.; Matteucci, M. D., *Nucleic Acids Res.*, **1983**, *11*, 8031-8036.
12. Potier, N.; Van Dorsselaer, A.; Cordier, Y.; Roch, O.; Bischoff, R., *Nucleic Acids Res.*, **1994**, *22*, 3895-3903.
13. Rozenski, J., Mongo Oligo Mass Calculator Tool [Online], **1999**, Available from: http://mods.rna.albany.edu/masspec/Mongo-Oligo , [Accessed 27/01/2014]
14. Integrated DNA Technolodies, IDT Biophysics UV Spectrum tool (percentage error: 4%), **2014**, Available from: http://biophysics.idtdna.com/UVSpectrum.html, [Accessed 27/01/2014] Parameters for the calculation found in the following publication: Tataurov, A.V.; You, Y.; Owczarzy, R., *Biophys. Chem.*, **2008**, *133*, 66-70.
15. Cantor, C. R.; Warshaw, M. M.; Shapiro, H., *Biopolymers*,**1970**, *9*, 1059-1077.
16. J. B. Birks, *Photophysics of Aromatic Molecules*, Wiley & Sons, **1970**.

Chapter 3: Photo-Switched Binding of Duplex DNA

3.1 Introduction

Following on from the successful synthesis of a series of double anthracene-tagged oligonucleotides, this chapter focuses on the development of methods to test their photochemical properties, as well as procedures to purify and characterise any resulting photoproducts.

This chapter will then progress into investigating techniques that can be used to analyse and quantify the effect of anthracene photodimerisation on the structure of an oligonucleotide. It can then be confirmed whether this leads to disruption in hydrogen bonding to the complementary sequence, which can potentially lead to the prevention of hybridisation.

As part of this research there will be a particular emphasis on the different variables that are available such as linker type and sequence composition, and how they can be used to optimise conditions for anthracene photodimerisation.

3.2 Photochemical Studies

3.2.1 Introduction to Molecular Photo-Physics and Photochemistry

When a chemical species absorbs a photon of light, an electron within that molecule become excited into different electronic states. In order to return to the ground state a number of photophysical processes can occur as described by the Jablonski diagram in Figure 1.

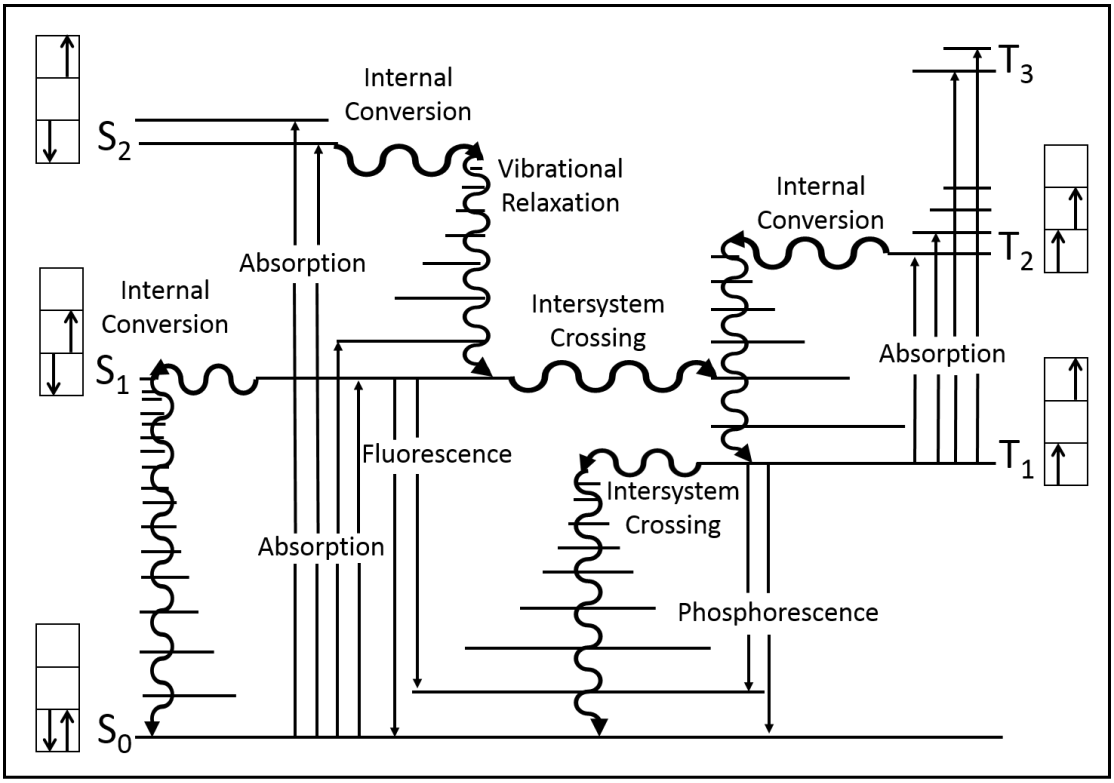
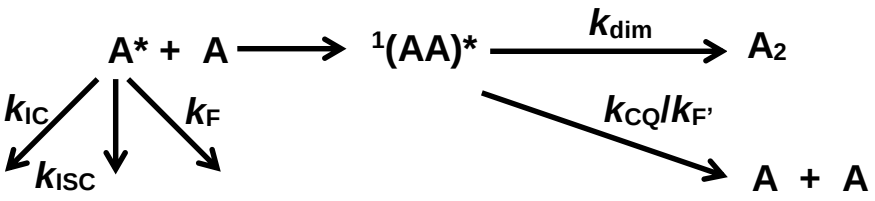


Figure 1. Jablonski diagram; Horizontal lines are vibrational energy levels, straight arrows depict radiative processes and curly arrows depict non-radiative processes

The electron is promoted from the singlet ground state S_0 into a higher vibrational energy level within the first or second excited state (S_1 or S_2) through absorption. Henceforth they rapidly relax to the ground state, either radiatively through fluorescence or non-radiatively through a combination of internal conversion and vibrational relaxation. Alternatively they can undergo intersystem crossing to the triplet states T_1 and T_2, returning to the ground state through a process of phosphorescence or another corresponding non-radiative process.[1]

Photochemical reactions are initiated when internal conversion and relaxation of an excited state leads to a ground state isomer of the initial molecule, alternatively they can occur when a molecule in the excited state undergoes an intermolecular addition to another reactant molecule in the ground state.

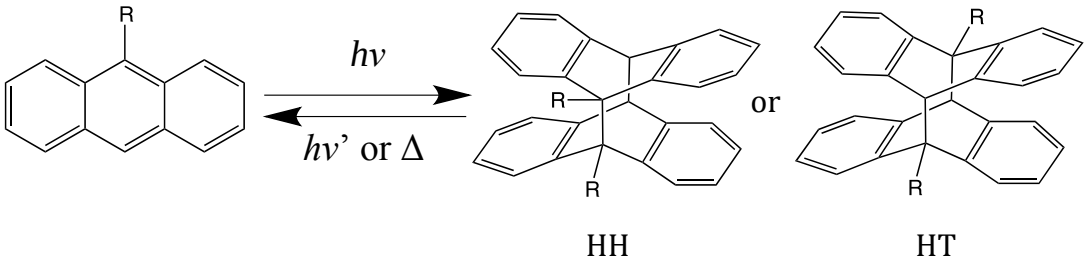
3.2.2 Anthracene Photodimerisation



Scheme 1. Kinetic scheme of the photodimerisation of anthracene (A). An excited state anthracene (A) can return to the ground state through a number of different pathways; internal conversion (IC), intersystem crossing (ISC) or fluorescence (F). Upon formation, an excimer (^{1}(AA)*) can undergo collisional quenching (CQ), excimer fluorescence (F') or dimerisation (dim).*

Anthracene is a classic example of a chemical species that can undergo photochemical reactions: It can photodimerise upon UV irradiation at 365 nm. The reaction is dependent on a photochemically excited anthracene (A*) coming into contact with a ground state anthracene (A) generating an excimer ^{1}(AA)*. An excimer is a short lived excited intermediate that is a prerequisite for photodimerisation. Upon formation it rapidly returns to the ground state through the formation of two new covalent bonds to produce a photodimer. But photodimerisation is only one of a number of competing pathways that can return anthracene to the ground state, as described by Scheme 1. The quantum yield of dimerisation (Φ_{dim}), a value which defines the number of photodimers produced per photon absorbed, is heavily dependent on the relationship between these

competing pathways. The photodimerisation of anthracene is a reversible process, with the dissociation reaction initiated upon the application of heat or UV irradiation < 300 nm. This process is therefore termed to be photochromic, (Scheme 2).[2]



Scheme 2. The reversible, photochemical dimerisation of anthracene, and the two possible outcomes of the dimerisation of substituted anthracene compounds.

Photodimerisation is a form of pericyclic reaction known as a cycloaddition that is only possible under photochemical conditions due to what is known as the Woodward-Hoffmann rules.[3] These state that the frontier molecular orbitals, i.e. the highest occupied molecular orbital (HOMO) of one component and the lowest unoccupied molecular orbital (LUMO) of a second component, must match in order to interact. This is important, as with this type of pericyclic reaction (a [4π + 4π] cycloaddition) the two new covalent bonds between the two anthracene molecules are generated in a concerted fashion. Therefore only the HOMO of an anthracene in the excited state can interact with the LUMO of a ground state anthracene, as it is only in this state that the frontier molecular orbitals are in the correct phase to generate both covalent bonds simultaneously, (Figure 2).[4]

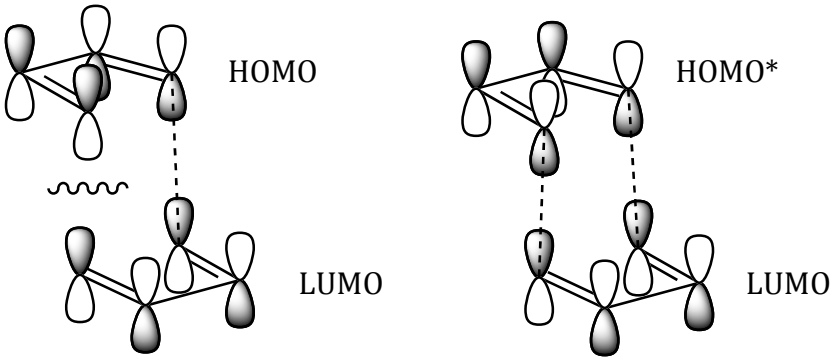


Figure 2. Frontier molecular orbital diagrams of the interaction between two anthracene molecules; only an anthracene in the photochemically excited state (*) can perform the [4π + 4π] cycloaddition reaction.

Figure 3A displays a simplified energy profile for the photodimerisation of anthracene that demonstrates the distance dependence (*d*) of the reaction. Anthracene in a singly excited state (S) reaches an energy minimum, at which point it either transfers into the double excited state (D) or returns to the ground state through excimer fluorescence ($h\upsilon_D$). The degree of overlap between these two states will determine the extent to which the excimer fluorescence and dimerisation pathways compete. Upon reaching the pericyclic minimum (P) there is a potential energy surface crossing to the ground state, where the

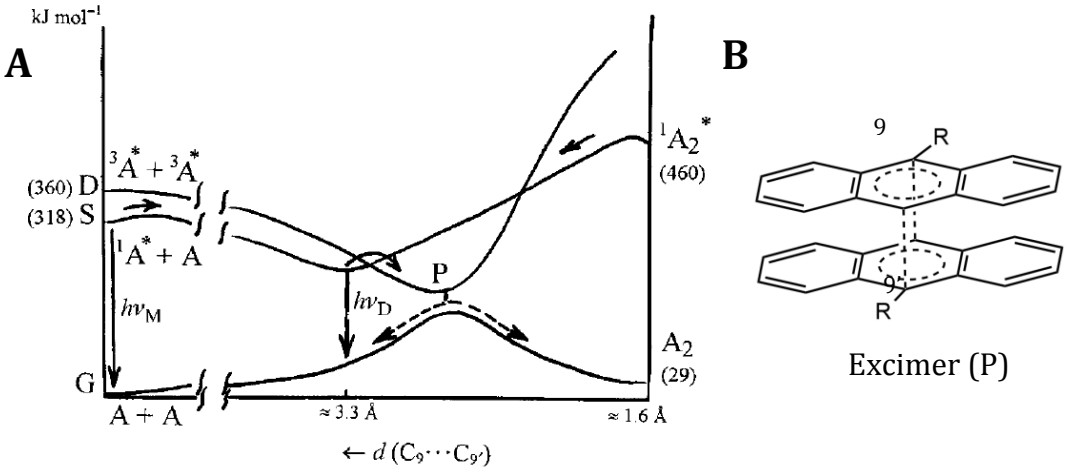


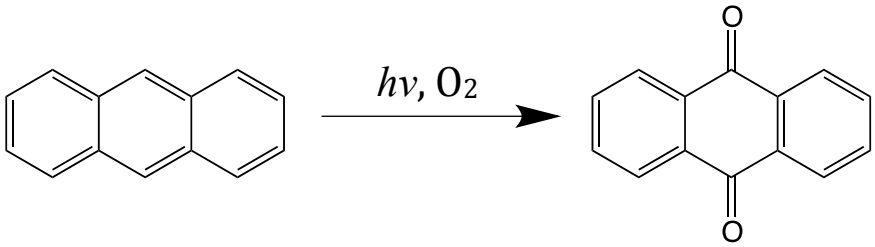
Figure 3. (A) Energy profile for the photodimerisation of anthracene, including the energy minimum at which the excimer (P) occurs. (B) Proposed structure of the excimer prior to photodimerisation. Image taken from ref. 2.

excimer, also known as a pericyclic intermediate (Figure 3B) will either dissociate to the monomeric anthracene (collisional quenching) or can overcome the energy barrier to generate the anthracene photodimer A_2.

Anthracene monomers substituted in the C9 position can photodimerise into two possible isomers, a 'head-to-tail' (HT) isomer, where substituent groups end up opposite each other, or a 'head-to-head' isomer, where substituent groups are on the same side (Scheme 2). The stability of each isomer is dependent on the substituent in each case, however the HH isomer is often the least stable due to steric clashes between the substituent groups.[5]

3.2.3 Method of Irradiation

The photodimerisation of anthracene is sensitive to the presence of oxygen, and once irradiated it can undergo oxidation and is converted to anthraquinone (Scheme 3).[6] To prevent this from occurring the sample needs to be degassed. In initial studies within this work this was done using a freeze-pump-thaw method, which involved placing a frozen sample under vacuum to remove the oxygen as it thawed. But it was subsequently found that it was just as effective to continuously degas the sample with an inert gas such as nitrogen or argon during irradiation.



Scheme 3. The photochemical oxidation of anthracene to anthraquinone.

To irradiate the sample, a 125 W mercury lamp is used as a light source (Figure 4), which is placed in a special housing fitted with a 365 nm (± 5 nm) interference bandpass filter (Edmund Optics, Part no. 65-675). This provides monochromatic light focused at the absorption wavelength of anthracene and will prevent damage to the oligonucleotide, which absorbs UV light at 260 nm. The lamp is also placed inside a water cooled jacket that prevents damage to the bulb, as well as heat transmitting to the sample, which could affect results. There was no observed increase in temperature within the sample during the course of the experiment.

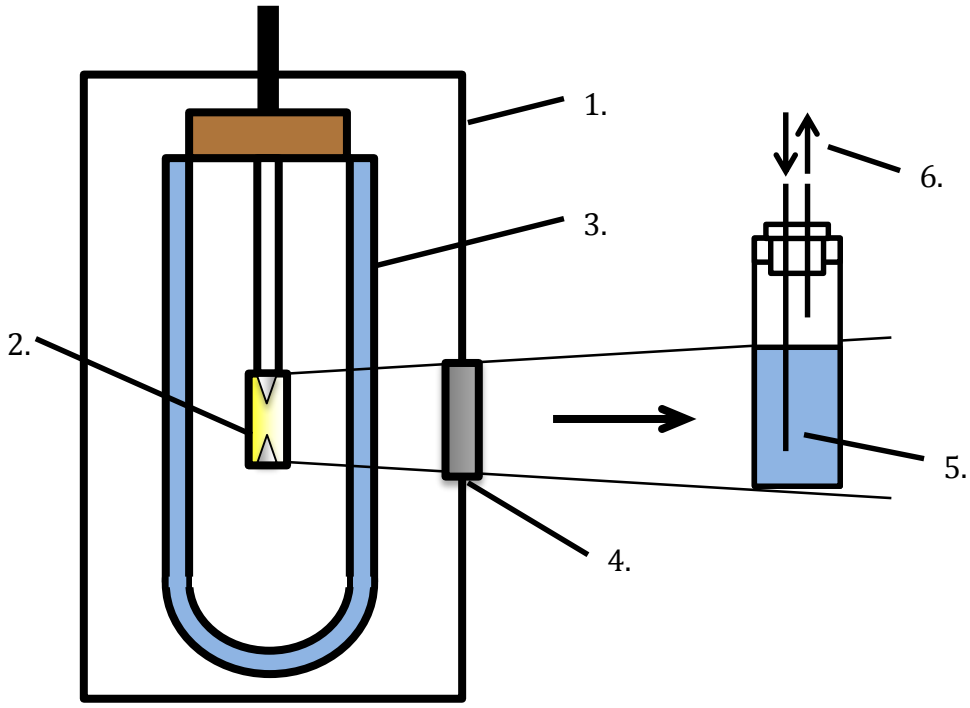


Figure 4. Schematic of irradiation setup including 1. Lamp housing, 2. Mercury lamp, 3. Water cooling jacket, 4. Bandpass filter, 5. Sample in quartz cuvette 6. Gas Flow.

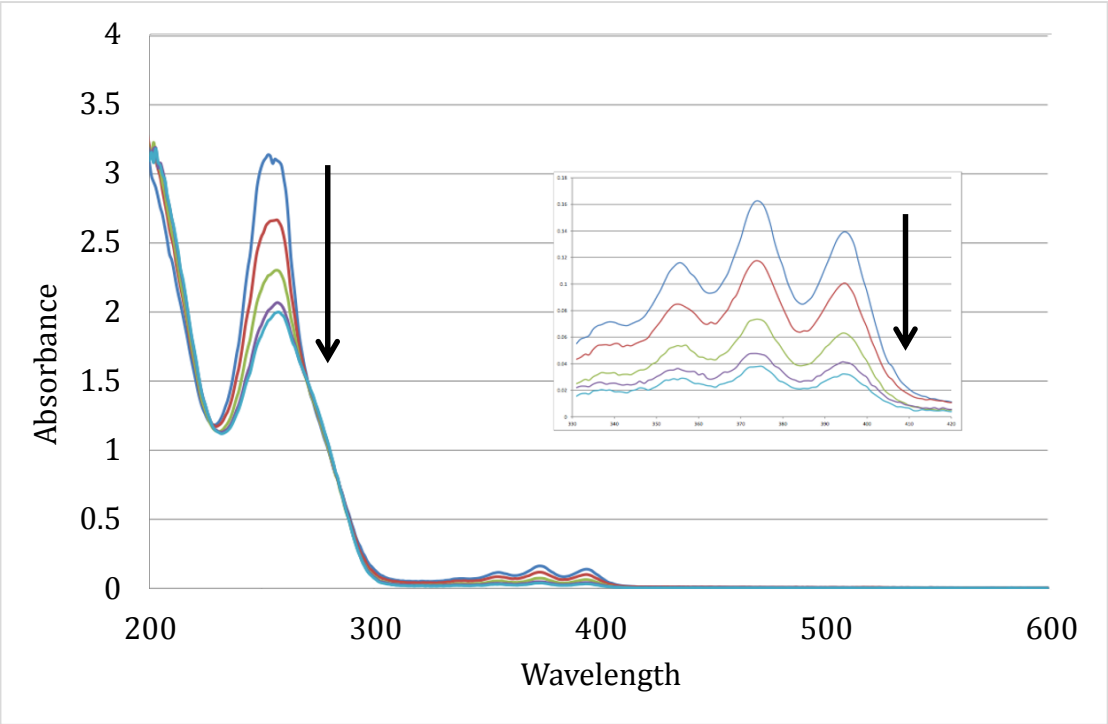


Figure 5. UV analysis of photodimerisation reaction recorded at set intervals

When two anthracenes are converted to a single photodimer upon irradiation there is a subsequent reduction in the π-conjugated system, with the product not absorbing UV light > 300 nm, as well as a significant decrease in absorption between 240 – 300 nm. This provides a useful way in which to monitor the reaction, as the sample can be irradiated within a quartz cuvette, which can be directly analysed at set intervals during the reaction using a UV spectrometer, producing spectra similar to those shown in figure 5.

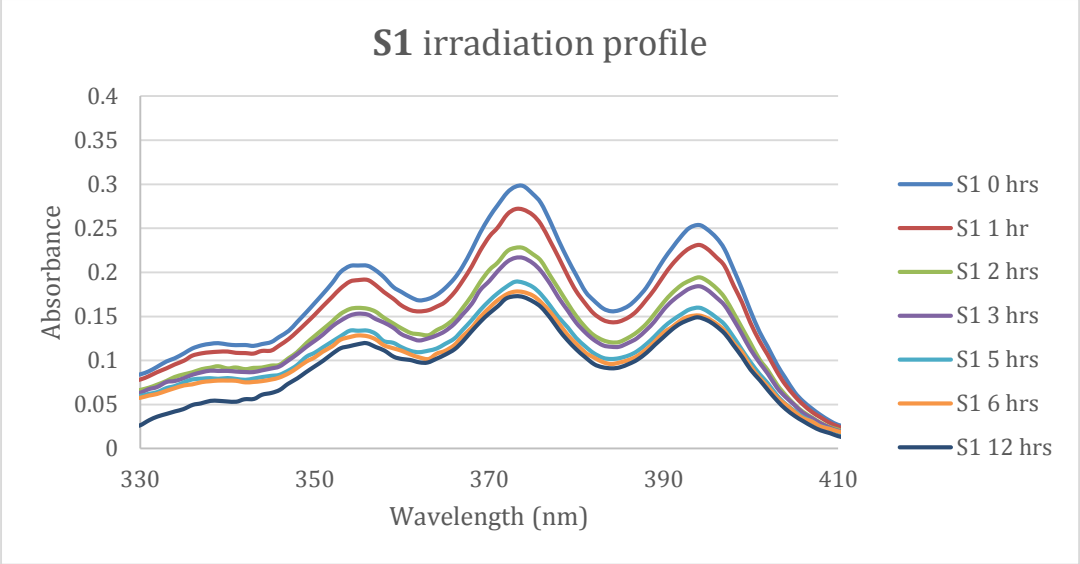


*Figure 6. Irradiation profile of sequence **S1**, the reaction progressed over a period of six hours, but the anthracene band remains after extended irradiation for twelve hours.*
S1: TGGACT**X**T**X**TCAATG

A buffered solution containing 20 μM of the sequence **S1** (100 mM NaCl, 10mM phosphate buffer pH 7.0) was irradiated at ambient temperature (21 °C) for a set interval of time and the progression of the reaction was monitored by UV-spectroscopy until no further decrease in the absorption was observed. The reaction progressed over a period of six hours (Figure 6), where the change in absorption gradually tailed off until the reaction was ceased. In repeated experiments with prolonged irradiation times, the anthracene band between 300 and 400 nm never fully diminished, however the significant drop in absorption signified a photochemical reaction and the samples were therefore submitted for HPLC analysis to analyse any photoproducts.

3.2.4 HPLC Analysis of Photoproducts

Figure 7 displays analytical HPLC chromatograms of samples collected at different stages of the reaction to show the change over time. As the reaction progressed two new peaks were observed in the HPLC chromatogram, each with a shorter retention time than that of the starting material (D_1 and D_2, Figure 8).

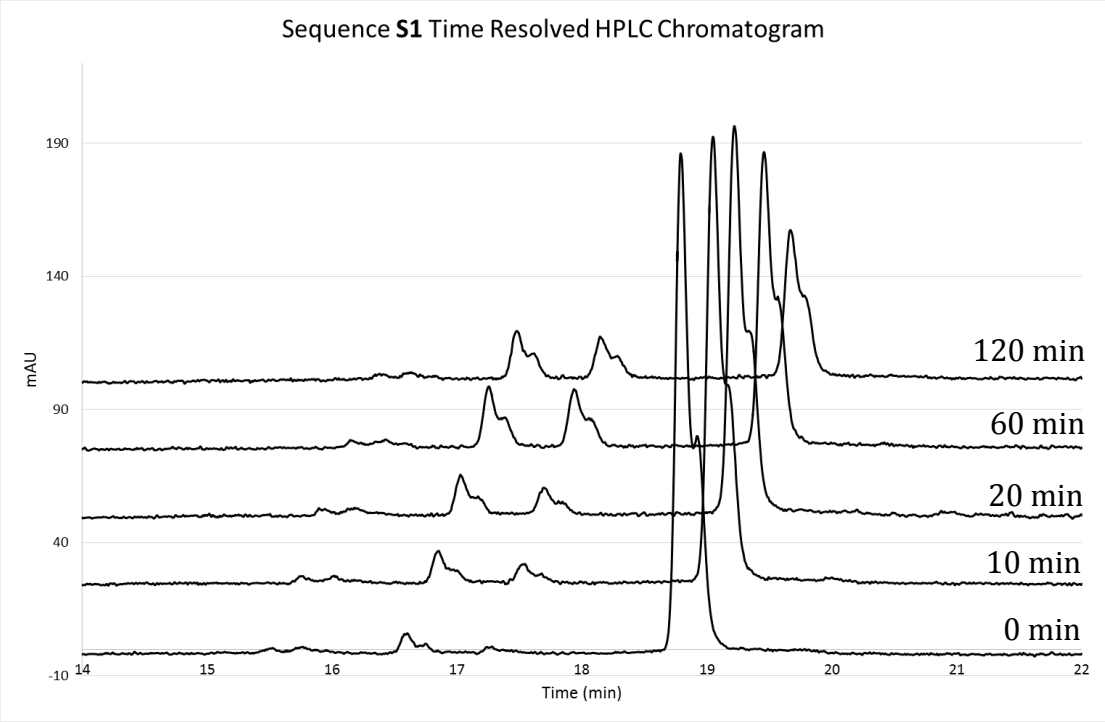


Figure 7. Overlaid HPLC chromatograms from samples taken at set intervals during the photochemical reaction to monitor the appearance of photoproducts.

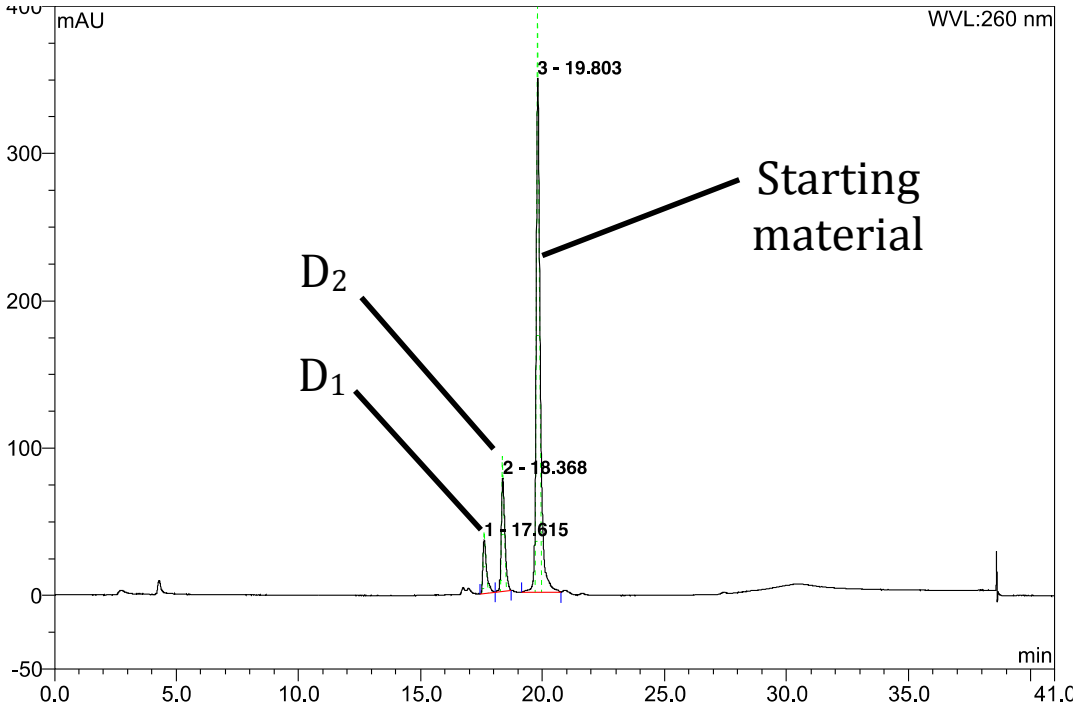


*Figure 8. The resulting HPLC chromatogram of sequence **S1** displaying the emergence of two new photoproduct peaks (D_1 & D_2).*

The photoproducts D_1 and D_2 were then isolated and submitted for electrospray mass spectrometric analysis, where each photoproduct gave a peak identical in mass to that of the starting material. This results is a good indication that the intramolecular anthracene photodimerisation reaction has taken place rather than any other intramolecular or intermolecular reactions. It is possible that the anthracene can photodimerise intermolecularly to a second **S1** oligonucleotide, however no additional peaks at double the mass of **S1** were observed in the mass spectrum and the probability of this occurring at the concentrations used during irradiation would be quite low. In addition to this anthracene could potentially react with neighbouring nucleobases during the irradiation process, however upon further testing with sequences using a single anthracene modification, as discussed in Section 3.9, no photoreactivity was observed indicating that this does not occur.

3.2.5 'Head to Head' vs. 'Head to Tail'

When two photoproducts were obtained from the photodimerisation reaction it was hypothesised that the 4D linker within this sample gave the anthracene enough degrees of freedom that it was possible to form both the HH and HT isomers of the photodimer.

To test this theory, purified photodimers were placed in a variable temperature UV spectrometer and heated to 60 ^{o}C, whilst the UV absorbance at 374 nm was monitored continuously through 1 minute intervals for a total of two hours. The results of this test indicated that the first photoproduct (D_1) was thermally more stable and remained relatively unchanged, whilst the second photoproduct (D_2) rapidly increased in absorption, signifying the thermal reversion to the starting material (Figure 9). Due to the fact that the HH dimer would be inherently the less stable of the two photodimers due to steric clashes between the substituent groups on the C9 position, it was assumed that the second photoproduct D_2 can be correctly identified as the HH isomer. The more stable HT photoproduct D_1 was therefore retained for further analysis, since it would be stable to VT-UV studies (Section 3.3.3).

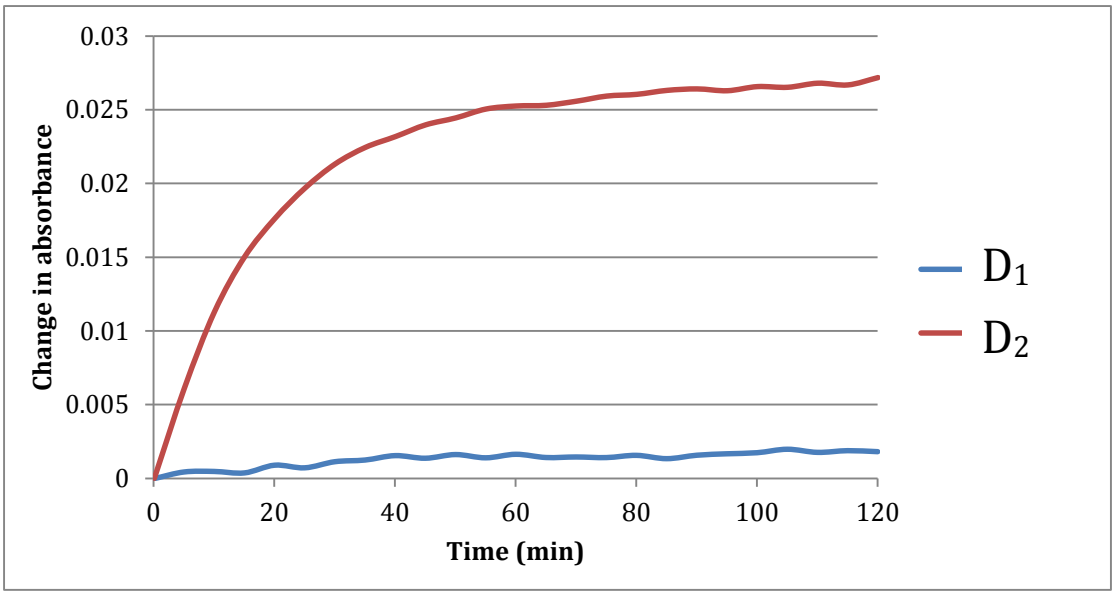
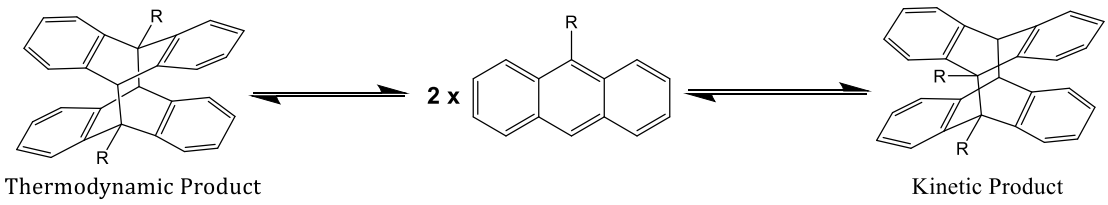


Figure 9. Thermal analysis of D_1 and D_2 photoproducts used to identify the corresponding isomers of the anthracene photodimer (60 °C).

With the two photoproducts now identified as different isomers of the anthracene photodimer, analysis of the HPLC chromatogram indicated that it is the less stable HH photoproduct that is produced in greater quantity. It would appear that the orientation of the two anthracene groups appended to the backbone of the DNA favour the formation of the HH dimer upon reaction, making it the kinetic product. However due to the apparent lack of stability in the resulting photoproduct, it is likely that the dimer reverts back to starting material under the ambient conditions used during irradiation. The rate of formation of the HT photoproduct is therefore much lower, however the greater stability of this isomer make it the thermodynamic product of the reaction (Scheme 4).



Scheme 4. The equilibrium between the two isomers of the anthracene photodimers, with the HT isomer as the thermodynamic product and the HH isomer as the kinetic product.

This observation may also explain why the photodimerisation reaction is not capable of reaching completion, even after extended irradiation times. At the beginning of the irradiation the rate of HH reversion is overshadowed by the forward dimerisation reaction and the drop in absorbance is greater. However as the reaction progresses and the rate of dimerisation diminishes an equilibrium is reached, also known as a photostationary state, where both processes are in balance. As the HT photodimer is the thermodynamic product and is stable at room temperature it is possible that given enough time, the ratio would shift to give 100% HT (D_1). However, the ratio between the starting material and the two photoproducts in the HPLC chromatogram would suggest that the rate of dimerisation of the HT isomer is relatively low, therefore trying to achieve 100% yield of the HT dimer under these conditions would take an indeterminate amount of time. Subsequent studies using a better optical filter resulted in much higher yields of the photodimer product, (Section 3.4.1).

3.3 Effect of Anthracene Linker Types

3.3.1 Experiment design

As discussed in Chapter 2, there are two features of the anthracene monomer that can be customised; the length of the alkyl spacer and the configuration of the stereogenic centres within the threoninol group, each of which will influence how the anthracene group is oriented once incorporated within an oligonucleotide. Whilst the length of the alkyl spacer will determine the overall reach and flexibility of the anthracene tag, the configuration of the stereogenic centres within the threoninol group will control how the tag is positioned with respect to the backbone of the oligonucleotide.

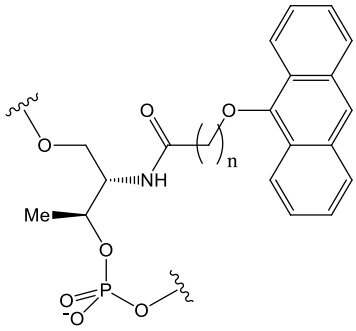


Figure 10. Anthracene monomer structure, including an adjustable alkyl spacer (n), and variable stereochemistry within the threoninol linker (shown here as the S,S configuration).

Earlier work by Duprey *et al.* attempted to investigate the relationship between the configuration of each linker and how it influences the orientation of the appended anthracene with respect to the DNA backbone. His research attempted to utilise the fluorescent response from anthracene to sense small changes in the local environment, which through a number of different strategies would allow the detection of single nucleotide polymorphisms (SNPs), (Section 1.8). In order to achieve the best results, Duprey created an array of sequences, each with a varied linker type, in order to find an optimal set of conditions to detect SNPs.[7]

For the work in this thesis it was considered important to investigate how the variation of linker types could affect the rate of the photodimerisation reaction and whether an optimal set of conditions could be identified, improving the yield of the photodimer. It was also expected that the variation of linker type would also have an effect on the extent of structural change that occurs upon photodimerisation; anthracene tags with shorter linkers that successfully dimerise may lead to a larger structural change than those with longer linkers in the same scenario, leading to decreased hybridisation.

Table 1. Sequences of the twelve DNA strands synthesised to study the effects of variation in linker type and configuration.

Set	Sequence Name	Linker Type	Sequence Composition (5'-3')
1	**S2**	4D	TGGACTC**X**C**X**CAATG
	S3	4L	
	S4	6D	
	S5	6L	
2	**S6**	4D	TGGACT**X**T**X**TCAATG
	S7	4L	
	S8	6D	
	S9	6L	
3	**S10**	4D	TGGAC**X**C**X**CTCAATG
	S11	4L	
	S12	6D	
	S13	6L	
-	**S0**	-	TGGACTCTCTCAATG
	T0		CATTGAGAGAGTCCA

To investigate these concepts, a series of sequences were synthesised with variations in both the linker length and linker configuration and were subjected to a series of irradiation tests. Subsequent HPLC analysis, as well as a full thermal melting analysis could then be carried out and the data cross analysed to see if an optimal set of conditions could be found. Table 1 lists the full set of sequences used in the following series of experiments. With knowledge gained from the experimentation of sequence **S1**, which only fully dimerised over the course of several hours, it was decided that the n = 4 linker would be used as the minimum linker length and would be compared to the longest linker length that had so far been synthesised, the n = 6. This would hopefully provide a reasonable reaction time, in addition to obtaining sufficient yields of the photodimer. For the same reason, the position of each tag was kept within one nucleotide with respect to each other to allow an adequate proximity of the anthracene groups.

This experiment will also investigate whether the position of the tags within the sequence has any effect on the results; these changes may not influence reaction rate or yield of photodimer, but the position of the 'kink' within the DNA sequence may have an effect on hybridisation when paired with the complementary strand. Therefore the sequences were classed into three sets, where the position of the anthracene tags was shifted in set increments from the 5' end towards the 3' within the chosen sequence. As a result of varying the placement of anthracene tags within the oligonucleotide sequence, the bases flanking each tag will also change. In sets 1 and 3 the anthracene tag are flanked by all G/C pairs, whereas in set 2 they are flanked by all A/T pairs. The difference in the number of overall hydrogen bonds within these flanking bases (3 hydrogen bonds for G/C vs. 2 hydrogen bonds for A/T) may also have an effect on how the stability of the duplex changes upon photodimerisation as any change in structure as a result would likely influence these base pairs the most.

3.3.2 Irradiation and Rates of Reaction

Each of the sequences described in Table 1 were prepared in a buffered solution (20 μM in 2 mL, 10 mM phosphate buffer, pH 7.0, 100 mM NaCl) and subjected to irradiation at 365 nm (± 5 nm). The absorbance at 374 nm was recorded at set intervals over a forty minute period to monitor the initial rate of reaction. Figure 11 displays the percentage change in absorbance over time, grouped within the three sets of sequences.

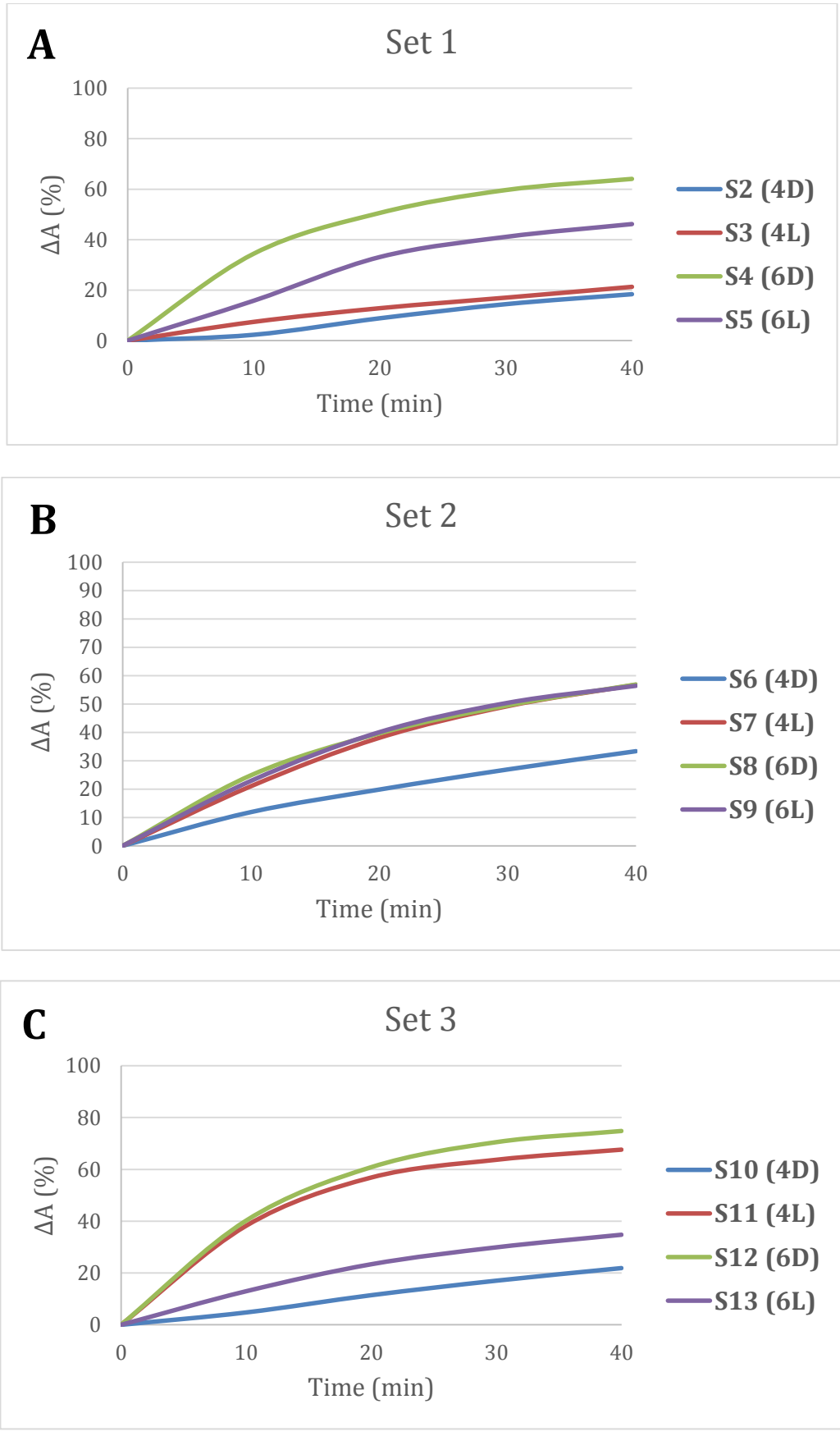


Figure 11 (A-C). Changes in the absorbance (365 nm) measured upon irradiation of anthracene modified sequences, giving an indication of sample reactivity. (20 μM samples irradiated in the single stranded state for 40 minutes, recording absorbance at 10 minute intervals)

Following irradiation, each sample was also submitted for analytical HPLC analysis (DMT On method). Table 2 displays the relative areas for both photoproduct peaks D_1 and D_2, as well as the peak of the unreacted starting material.

Table 2. Relative area of peaks observed in HPLC analysis immediately after irradiation. The D_2 photodimer is the principle photoproduct in most cases.

Set	**Sequence**	**HPLC Rel. Area (%)**			
		D_1	D_2	Ratio	Starting Material
1	S2 (4D)	3.1	7.6	0.41	89.3
	S3 (4L)	2.6	8.4	0.31	88
	S4 (6D)	16.2	26.3	0.62	57.5
	S5 (6L)	7.7	10.6	0.73	81.7
2	S6 (4D)	7.5	13.6	0.55	78.9
	S7 (4L)	9.7	27.2	0.36	63.1
	S8 (6D)	21.1	19.0	1.11	60
	S9 (6L)	16.3	18.8	0.87	64.9
3	S10 (4D)	2.7	9.4	0.29	87.9
	S11 (4L)	9	52.7	0.17	38.2
	S12 (6D)	10.9	37.9	0.29	51.2
	S13 (6L)	4.9	12	0.40	83.1

Upon analysing the UV data it was apparent that in most cases the n = 6 linker had the highest initial reactivity, with the 6D sequence outperforming others in both sets 1 and 3. Whilst there were exceptions to this trend, it would appear that the longer linker has improved reactivity as it gives the anthracene tags more flexibility and reach, therefore providing an increased chance of photodimer formation. When looking at the HPLC data it was again apparent that the 6D linker gave the best results, with higher yields of the more thermally stable D_1 photoproduct than any other sequence within the given set. Sequence **S8** also produced a higher ratio of the D_1 photodimer in comparison to D_2, unlike any other sequence tested so far.

From looking at earlier molecular modelling research by Duprey, it is known that the D configuration of threoninol orientates the anthracene groups away from the bulky nucleobases within the single stranded oligonucleotide.[8] This feature may help improve the probability of collision and formation of the photodimer, but with the longer linker lengths used in this study which provide a high degree of flexibility, both D and L configurations would be capable of orientating themselves within a position so that dimerisation can take place. Therefore the threoninol configuration is not such an important factor in this case, which would explain why there was no overall trend observed in terms of reactivity.

3.3.3 Melting Point (T_m) Experiments

To evaluate the effect anthracene modification and subsequent photodimerisation reactions had on nucleic acid structures, the duplexes were analysed through a series of thermal melting point experiments. The experiment uses a variable temperature UV-spectrometer (Varian, Cary 5000), which gradually ramps the temperature at a rate of 0.5 °C/min over a range of 15-85 °C and records the absorbance at 260 nm at 0.5 °C intervals. Due to the effect of hyperchromicity, the change in absorbance plots a sigmoidal curve as the DNA dissociates or anneals. Three successive ramps are recorded, two dissociative and one annealing, and the spectrometer software generates a T_m value through derivative calculations for each ramp. The T_m values for the two dissociative ramps are averaged to give the final T_m for each sample.

The rate at which the temperature is ramped in a melting point experiment is important as slower increases in temperature provide more accurate data, due to internal temperature heterogeneity. Therefore the ramp speed of DNA melting experiments are typically set between 0.1 and 1 °C/min to provide the most accurate results. All T_m data in this study was collected with a ramp speed of 0.5 °C/min, which combined with the collection of a second dissociative ramp should provide a sufficient level of accuracy. The experiment could be set up to collect more dissociative ramps, which would improve accuracy further, however due to the sensitivity of the anthracene photodimer towards extended heating the experiment was restricted to two ramps.

The removal of a nucleobase within an oligonucleotide reduces the number of effective hydrogen bonds that can be formed with the complementary strand, reducing the overall stability. In this instance two nucleobases are removed from the sequence to be replaced with anthracene tags. Therefore to investigate whether the modified sequences were still capable of forming stable duplexes with the complementary strand **T0,** the T_m data for each sequence was collected prior to irradiation, (Table 3). In comparison to the T_m data collected for the unmodified duplex **S0**/**T0** (61 °C), all samples were observed to have a reduced thermal stability due to the anthracene modification, but were reasonable enough to suggest that each sequence would generate a stable duplex at room temperature.

*Table 3. T_m analysis of the D_1 photodimer for sequences **S2** to **S13** paired with **T0**. Samples were analysed in the range of 15-85 °C (5 μM sample, 10 mM phosphate buffer, pH 7, 250 mM NaCl). The table provides T_m data before (SM) and after (D_1) irradiation, and hence the drop in T_m (Δ)* as a result of photodimerisation for each sequence.

Set 1: TGGACTC**X**C**X**CAATG
Set 2: TGGACT**X**T**X**TCAATG
Set 3: TGGAC**X**C**X**CTCAATG

		Set 1 (°C)	**Set 2** (°C)	**Set 3** (°C)
4D	T_m (SM)	48.5 (±0.7)	43 (±0)	48 (±0)
	T_m (D_1)	41.5 (±0.7)	35.5 (±0.7)	38 (±2.8)
	Δ	-7	-7.5	-10
4L	T_m (SM)	45.5 (±0.7)	39.5 (±0.7)	45 (±0)
	T_m (D_1)	36.5 (±2.1)	32.5 (±0.7)	38.5 (±0.7)
	Δ	-9	-7	-6.5
6D	T_m (SM)	44 (±0)	37.5 (±0.7)	44.5 (±0.7)
	T_m (D_1)	39 (±1.4)	37 (±0)	38 (±1.4)
	Δ	-5	-0.5	-6.5
6L	T_m (SM)	43.5 (±0.7)	37.5 (±0.7)	41.5 (±0.7)
	T_m (D_1)	39.5 (±0.7)	34.5 (±0.7)	37 (±1.4)
	Δ	-4	-3	-4.5

To measure the disruption in hybridization effected by the photodimerisation reaction, the D_1 photoproduct of each sequence was isolated through preparative HPLC and subsequently analysed by thermal melting (T_m) analysis. Table 3 lists the subsequent change in T_m (Δ) effected by photodimerisation upon pairing the isolated D_1 photodimer with the complementary strand **T0**. As can be seen from the data, photodimerisation of the anthracene tags leads to a drop in the thermal stability of the duplex when paired with the complementary strand, ranging from a 0.5 to 10 °C decrease in the T_m. Upon photodimerisation it is proposed that the change in structure and reduction of flexibility of the modified strand will reduce

the number of effective hydrogen bond and π-stacking interactions that can occur when paired with the complementary sequence. This change will be observed in a reduction in the T_m value, or as seen in Figure 12, a shift in the sigmoidal curve to a lower temperature range.

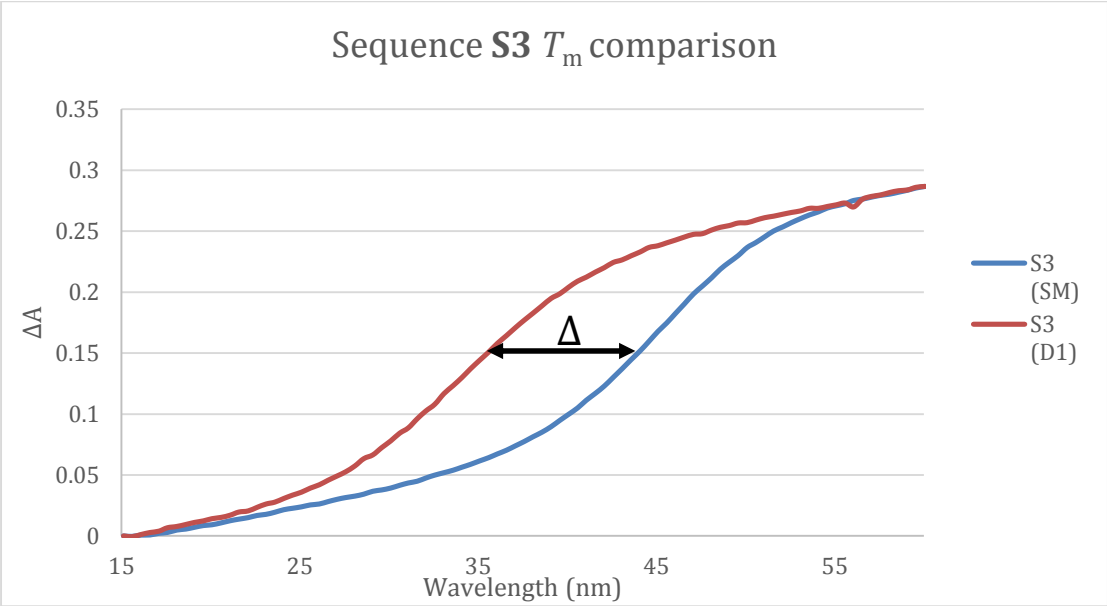


*Figure 12. T_m analysis before and after irradiation displaying the resulting drop in thermal stability of the **S3-D1/T0** duplex (Δ).*

Notably it is the shorter n = 4 linker that overall provides the most significant change in duplex stability, and the T_m for each sample is several degrees lower than its longer linker counterpart. The n = 4 linker, with fewer degrees of freedom, will require the structure of the oligonucleotide to become more significantly altered in order to accommodate the anthracene photodimer, which would explain the improved results for these sequences. This also correlates with earlier irradiation studies, suggesting that the greater distance between anthracene tags reduces the probability of collision.

When looking for trends in terms of variation in linker configuration, in general the best results were obtained for sequences that use the S,S (D) isomer of threoninol. For example, the highest values for Δ observed for both the *n* = 4 (-10) and *n* = 6 (-6.5) alkyl spacer involve the use of a S,S (D) threoninol linker. In this configuration, the single stranded DNA must either become more structurally rigid or altered in order to accommodate the photodimer, hence it is more difficult to achieve hybridisation and the value for Δ is greater.

When comparing data between the three sets of sequences it is important to note that due to the way in which the sequence is modified with anthracene sets 1 and 3 have additional hydrogen bonding due to a greater number of G/C base pairs. This is reflected in their T_m data when compared to set 2, displaying greater stability in all cases. Therefore when the anthracene in each sequence is photodimerised it is clear that this has a greater effect for samples in sets 1 and 2 as it is capable of disrupting more hydrogen bonding within the neighbouring G/C base pairs. In most cases this only provided minor improvements to the Δ value, but it was still important that all variables were monitored to see how they affect the outcome, so that a full optimisation could be achieved.

3.4 Base Variation Studies

With a better understanding of the significance of the linker length and its influence on the thermal stability of duplex DNA structures, a new set of sequences were developed that fixed the linker type and instead varied the number of bases that separated each anthracene tag. It was hypothesised that increasing the distance between each tag would induce larger structural changes upon

photodimerisation, leading to reduced hydrogen bonding to the complementary sequence. However by increasing the distance between tags there will be a reduction in the efficiency of the photodimerisation reaction. Therefore the linker type chosen for these sequences used the longer n = 6 alkyl spacer, giving the anthracene the maximum degree of mobility. When looking at the earlier T_m results displayed in Table 3, the n = 6 linkers gave poorer results in terms of ΔT_m for the single base separation sequence they were tested on, but gave the best results in terms of reactivity and yield of photodimer. Therefore this linker may be better suited to sequences with larger base separations, as the greater reach they give to the appended anthracene would potentially provide both sufficient reactivity and the larger structural changes required to achieve improved values for Δ upon dimerisation. Table 4 lists the three sequences synthesised in order to test variations in base separation. Based on information gathered from the previous study, the 6D anthracene unit was placed centrally within the sequence, with an increasing base separation of one, three and five base pairs between each tag.

*Table 4. Sequences of the three strands synthesised to study the effects of varying base separation between anthracene tags, as well as the three resulting photodimers (**X** = anthracene).*

Sequence Name	Sequence Composition (5'-3')
S14	TGGACT**X**T**X**TCAATG
S15	TGGAC**X**CTC**X**CAATG
S16	TGGA**X**TCTCT**X**AATG
S14-D$_1$	TGGACT**X**T**X**TCAATG
S15-D$_1$	TGGAC**X**CTC**X**CAATG
S16-D$_1$	TGGA**X**TCTCT**X**AATG

3.4.1 Filter Testing and Irradiation

In order to achieve the best results with these newly synthesised sequences, which could potentially have a lower reactivity than their predecessors, an investigation was carried out focusing on improving the irradiation set-up to give the best photochemical conditions possible. This involved improving the lamp housing with increased shielding, better lamp and filter alignment and deep cleaning of the reaction cuvettes to the best optical standards. But what proved to be most crucial was testing the integrity of the UV bandpass filter by measuring the percentage transmittance using a UV spectrometer. It was observed that the filter being used had a very poor transmittance at the 365 nm wavelength that it was designed for and was not uniform across the entire surface of the filter. This may be as a result of thermal degradation that has occurred during the lifetime of its usage.

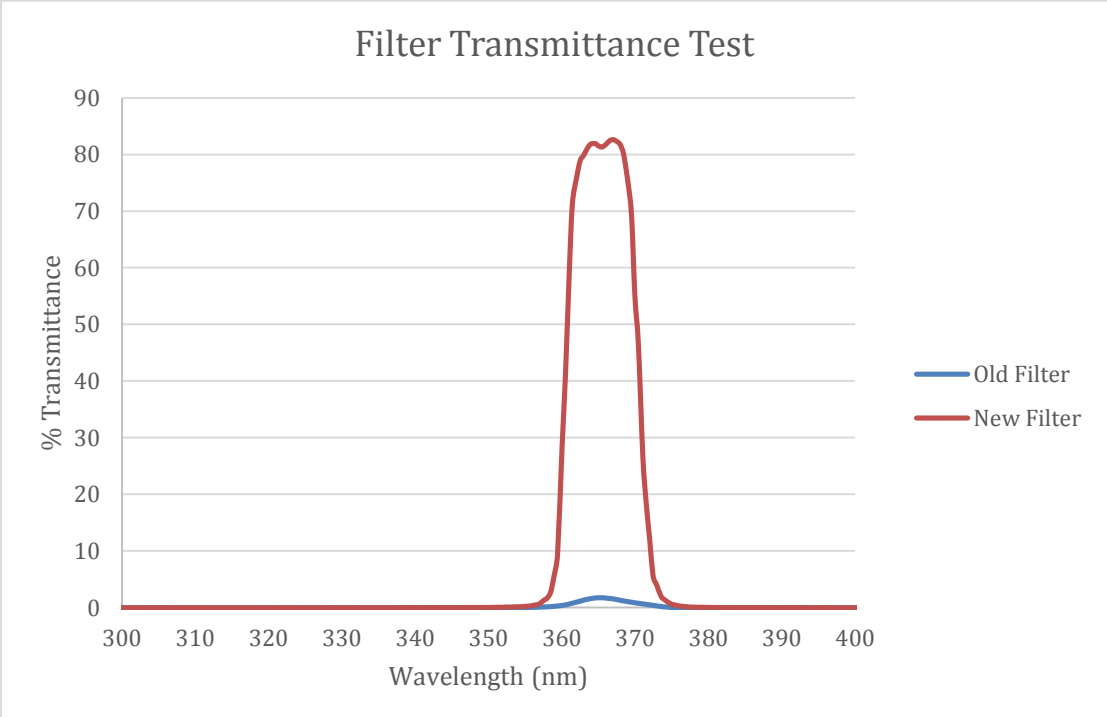


Figure 13. Transmission spectrum displaying the analysis of the old and new bandpass filters. The new filter allows a much greater transmittance of the desired 365 nm UV light, which should improve photochemical reactions.

A replacement filter was purchased with similar specifications to the existing one but had been manufactured with improved thermal resistance (Edmund Optics). When the transmittance was tested in comparison with the previous filter, there was a stark difference in the amount of light that was transmitted, with a well-defined band centred around the excitation wavelength, (Figure 13). In the light of these results, the new filter was adapted and fitted to the lamp housing to be used on future samples. Further to this the filter was regularly checked for wear so that future experiments were not compromised.

With the irradiation set-up fully optimised, the new batch of sequences was tested to see if the new filter made a significant difference to the dimerisation reactivity. What was observed was a marked improvement on earlier experiments, with reaction times reduced from over six hours to under forty minutes compared with the old filter. This result was also observed for sequence **S16,** which has five base pairs separating each anthracene tag, indicating that the increased distance was not as much a limiting factor towards the reactivity as previously thought. The increased light irradiation from the lamp source has greatly improved the efficiency of anthracene excitation, leading to an increased chance of photodimerisation.

As before the anthracene band did not fully disappear from the UV spectrum, indicating that even with this improved set up an equilibrium point is reached during the reaction, preventing full conversion to the photodimer. This was confirmed upon HPLC analysis as a starting peak still appears even after extensive

irradiation (Figure 14). The more significant observation within the HPLC analysis is the ratio between each of the peaks in the chromatogram, with yields of the D_1 photoproduct greatly improved compared with previous results. With this improved set up it is clear that there are a greater amount of photons reaching the sample, which increase the probability of an excited state anthracene reacting with a ground state anthracene in the correct orientation to produce the more stable HT photoproduct. In relation to the discussion in Section 3.2.5, this has had the effect of shifting the equilibrium to the thermodynamic product.

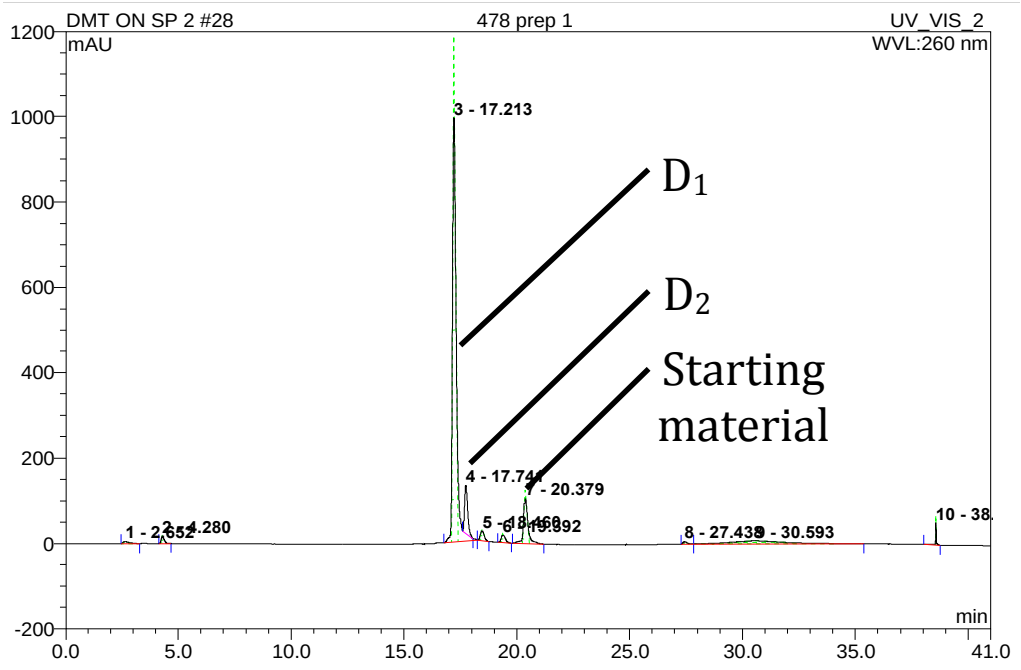


*Figure 14. HPLC analysis of sequence **S16** when irradiated using the new filter showing a significant shift to the D_1 photoproduct.*

3.4.2 T_m Analysis

*Table 5. T_m data for sequences **S14** to **S16** with the complementary **T0** (5 μM sample, 10 mM phosphate buffer, pH 7.0, 250 mM NaCl). There is increased disruption to hydrogen bonding when the base separation is increased, sequence **S16** displays no T_m transition within the range of detection (5 - 85 °C).*

	S14	**S15**	**S16**
SM	37.5	45	35
D_1	37	25	< 5
Δ	-0.5	-20	> -30

T_m analysis of the isolated D_1 photoproducts for sequences **S14**-**S16** revealed that a striking trend was apparent when compared with the corresponding starting materials, (Table 5). The larger the number of base pairs between each anthracene tag, the greater the reduction in the observed T_m between the photodimer and the complementary strand. Sequence **S14** has a single base separation, and as before displays little difference compared with the starting material, but increasing the number of bases to three and five in **S15-D_1** and **S16-D_1** respectively leads to a significant disruption in duplex stability. This effect occurs to such an extent that **S16-D_1** displayed no T_m transition at all within the temperature range of the melting point experiment. The test was repeated using altered settings to increase the range to as low as 5 °C and again no transition was observed, indicating that no duplex formation was possible under these conditions (Figure 15). Therefore whilst most samples tested so far may reduce the total number of hybridised oligonucleotides at room temperature, only sequence **S16** displays the desirable ON/OFF binding that could prove useful for further applications.[9]

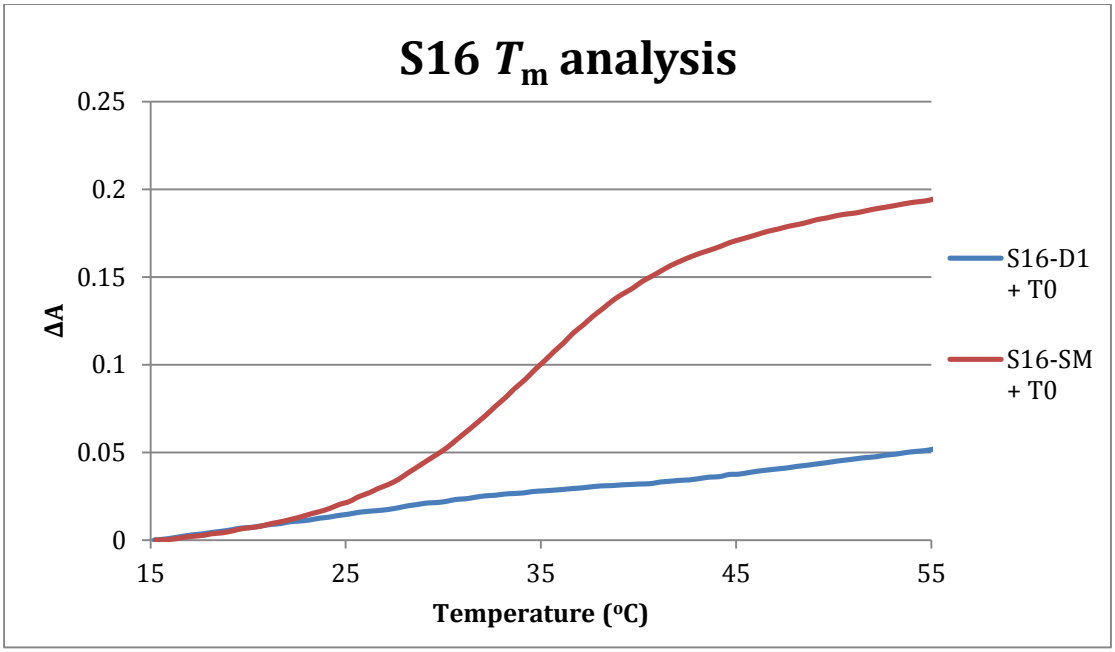


*Figure 15. Sequence **S16** displays an ON/OFF behaviour upon photodimerisation, with no T_m transition observed between the D_1 photoproduct and **T0** complementary sequence.*
***S16**:* TGGAXTCTCTXAATG
***T0**:* ACCTGAGAGAGTTAC

3.5 Gel Electrophoresis

3.5.1 Introduction to Gel Electrophoresis

Gel electrophoresis is a well-established technique that uses a porous gel material such as agarose or polyacrylamide to separate charged macromolecules by applying an electric field. Samples migrate through the gel through an electromotive force and travel at different speeds depending on their charge to size ratio, therefore the technique is ideal for use in the separation of biomolecules such as DNA, RNA and proteins.[10,11]

Samples are loaded in wells within the gel and a current is applied to initiate the separation as components of each sample migrate towards the opposite electrode. When the separation is complete the gel slab is removed from the apparatus and

is imaged through a number of techniques such as staining or UV transillumination, as the nucleic acid fragments will be invisible to the naked eye.

Common applications of gel electrophoresis such as the Sanger sequencing procedure, which is used to determine DNA sequences, separate nucleic acid fragments depending on the number of nucleotides within each sequence.[12] The fragments within each sample are single stranded, but will tend to fold into complex shapes or will partially hybridize making them harder to separate. In order to improve resolution, denaturing agents such as urea or formamide are added to the gel to disrupt any hydrogen bonding interactions, keeping the fragments in a rod-like shape. A large voltage is also applied, which generates heat, again preventing the formation of any secondary structure. These conditions, known as denaturing conditions, generate defined bands that are well separated within the gel image making it easier to identify each fragment within the sample.[13]

Alternatively, gel electrophoresis can be run using what is known as 'native' conditions, which promote hydrogen bonding and the formation of secondary structure. Under these conditions samples which are able to form a DNA double helix migrate slower than those which remain single stranded; this setup is therefore ideal for providing a visual confirmation that the photodimerisation has prevented hybridisation to the complementary strand. As described below, a number of experiments were set up using these conditions to investigate whether gel electrophoresis can be used to confirm the results from T_m tests that the

formation of secondary structure is limited or completely inhibited by the dimerisation of the anthracene tags.[10,11,14]

3.5.2 Native Gel Electrophoresis Procedure

The native gel electrophoresis experiments were carried out using a vertical slab gel electrophoresis kit (Hoefer SE400) with a 20 % polyacrylamide gel cast with a 1.5 mm thickness. Samples were dispensed into eppendorfs and lypholised using a rotary speed vac, before being re-dissolved in 20 μL of the running buffer. The samples were then heated to 80 ^{o}C for five minutes and slow cooled to room temperature to aid annealing of complementary strands and 10 μL of glycerol was added, which is used to improve gel loading.

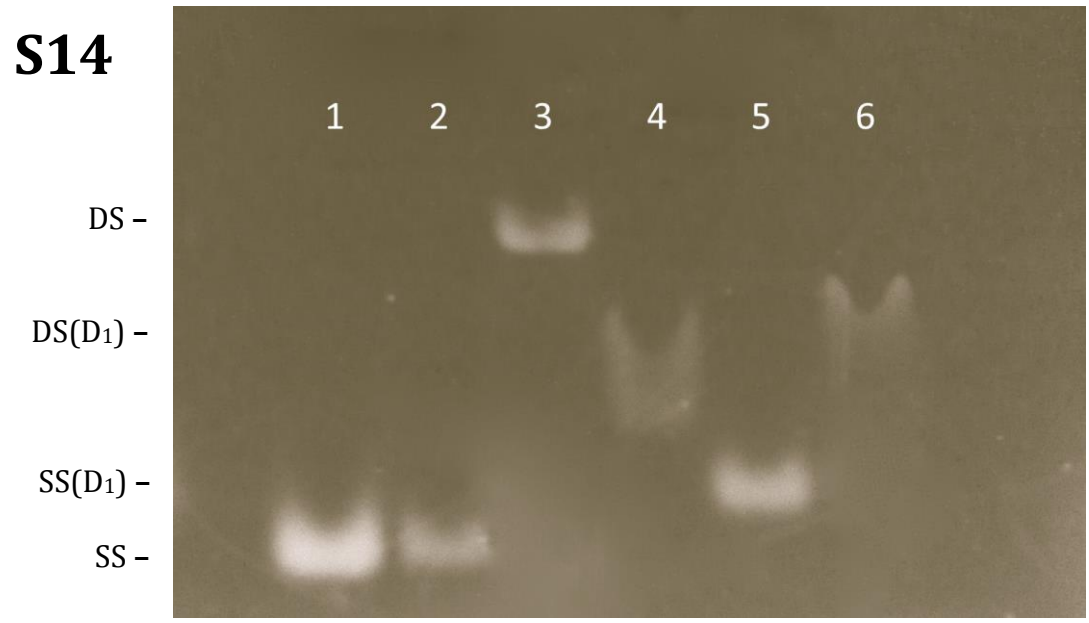
The most rapid and convenient method of visualisation in this experiment is through what is known as UV shadow transillumination. A completed gel slab is placed on top of a TLC plate card and then positioned under a UV lamp. DNA bands within the gel absorb the UV light, casting a shadow onto the TLC plate, which reflects the UV light but now within the visible spectrum allowing the viewer to see the shadow created by each band.[15] In order for this method of visualisation to be used each sample must be concentrated enough to produce a visible shadow, therefore each sample to be run was measured to a minimum of 0.5 to 1 OD.

Under the native conditions used in this experiment the normal tracking dyes bromophenol blue and xylene cyanol can be used to monitor the progression of DNA migration. However it should be noted that single stranded bands travel ~10-

15% faster than the leading bromophenol blue dye, therefore experiments must be stopped at a suitable time in order to prevent samples running off the gel slab.

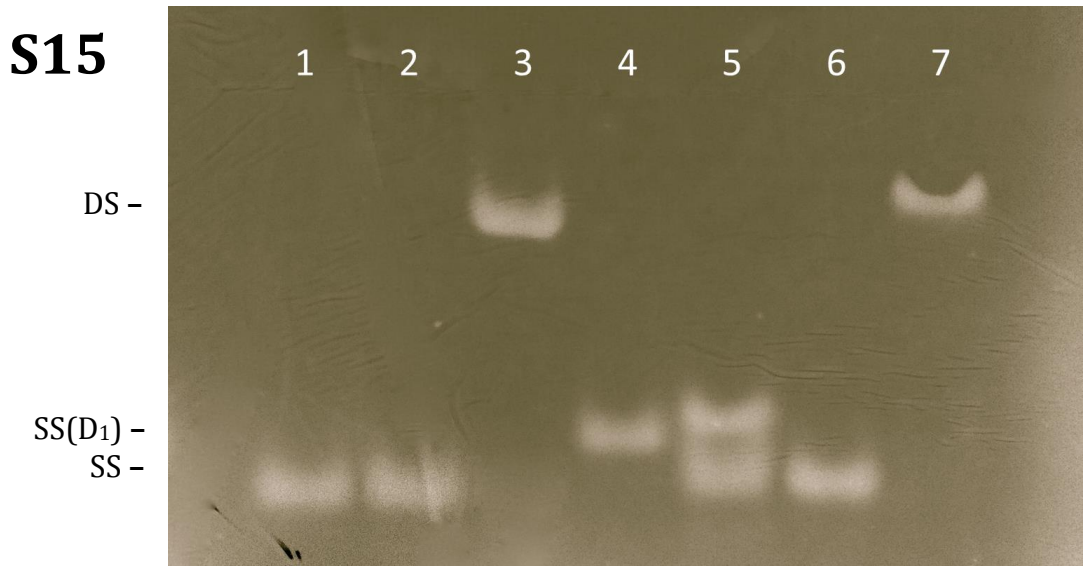
3.5.3 Gel Electrophoresis Results

The following set of images depict the gel electrophoresis results of sequences **S14**-**S16** after separation under native conditions. Each sample was run before and after irradiation and tested with and without the presence of the complementary strand. As a reference, the unmodified sequence was run in parallel in order to make a good comparison and evaluate how anthracene modification and subsequent dimerisation affected duplex binding.

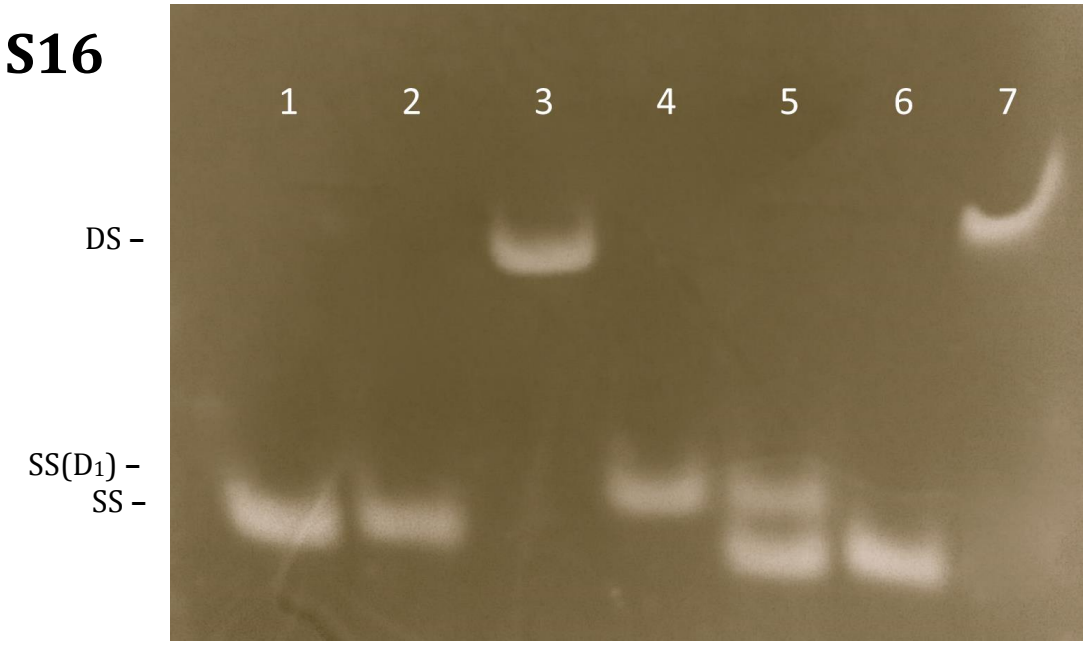


Lane	1	2	3	4	5	6
Sample	**S0**	**T0**	**S0/T0**	**S14-D1 /T0**	**S14**	**S14/T0**

Figure 16. Figure continued on next page.



Lane	1	2	3	4	5	6	7
Sample	**S0**	**T0**	**S0/T0**	**S15-D1**	**S15-D1/T0**	**S15**	**S15/T0**



Lane	1	2	3	4	5	6	7
Sample	**S0**	**T0**	**S0/T0**	**S16-D1**	**S16-D1/T0**	**S16**	**S16/T0**

*Figure 16. (**S14** – **S16**), UV shadow images taken after native gel electrophoresis analysis. Samples migrate at different speeds depending on the presence of secondary structure. These images can be used to confirm a samples ability to hybridise after photodimerisation of the appended anthracene groups. (20% native PAGE gel, 50 mM NaCl, 1 x TB buffer, 20 hour run time, 100V)*

As can be seen from the UV shadow images, there is a clear discrimination between single (SS) and double stranded (DS) DNA, with the secondary structure of the double stranded samples significantly reducing the speed of migration. In the images for each sequence the starting material sample successfully binds to the complementary strand prior to irradiation (lanes 6 & 7) and shifts to the upper band position in a comparative manner to the control sequences **S0**/**T0** (lanes 1, 2 & 3).

Post irradiation, the isolated D_1 photodimer (lane 4) was observed to have an altered migration speed compared with the unmodified sample **S0** (lane 1) and the sample prior to irradiation (lane 6). Under native conditions single stranded DNA is 'rod-like' in shape, with a lower cross-sectional area and can therefore pass through the gel easily. This result would indicate that the structural alteration that is imposed by the photodimer has increased the cross-sectional area and therefore proceeds a lower speed through the gel. When samples **S15-D$_1$** and **S16-D$_1$** were paired with the complementary strand (lane 5) there was no binding observed, with each of the samples migrating at their respective speeds, resulting in two bands on the gel. As for sample **S14-D$_1$**, which displayed the lowest value for Δ upon dimerisation, the photodimer (lane 4) can still achieve binding to the complementary strand and therefore remains unchanged compared to the starting material (lane 6).

The DNA staining agent ethidium bromide (EtBr) is commonly used as a fluorescent tag to detect DNA on a gel slab through transillumination techniques.[16] It works by binding to the DNA through intercalation, which increases the

compounds fluorescence ten-fold, making it possible to be detected even with very low amounts of DNA present. Since EtBr functions through intercalation it preferentially binds to double stranded DNA where binding is stronger due to the more regular structure; it is therefore primarily used in biological studies where DNA is in the double stranded form such as PCR reactions.[17] In this study the preferential binding of EtBr can be used to an advantage, because when the gel slab was stained with EtBr and imaged through transilluminescence (Figure 17); only bands in the upper part of the gel with a slower migration speed appear in the image. This confirms that only these bands contain DNA in the double stranded form, and that for sequences **S15-D$_1$** and **S16-D$_1$** no hybridisation occurred between the photodimer and complementary strands since there was no band visible in lane 5, (Figure 17).

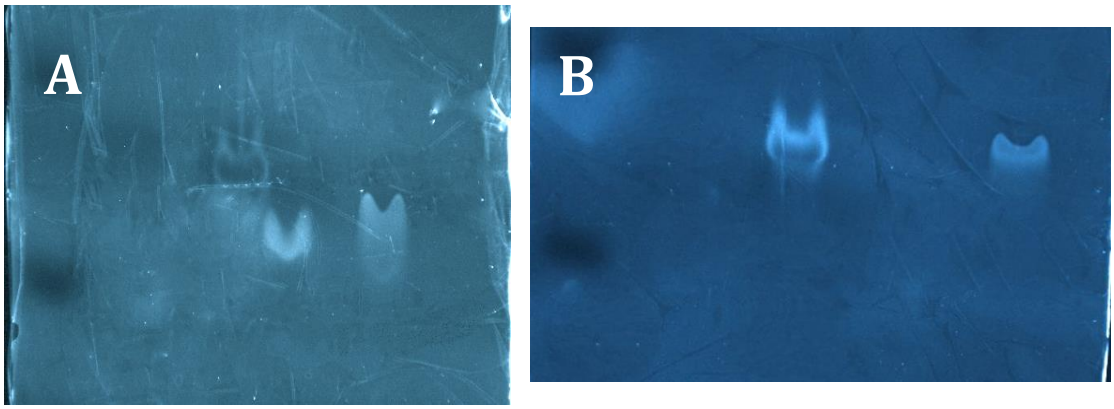


Figure 17. Ethidium bromide staining images. (A) ***S14****, no destabilisation occurs upon photodimerisation, therefore three bands appear for* ***S0/T0****,* ***S14-D$_1$/T0*** *and* ***S14/T0****. (B).* ***S15****, only two bands appear, as there is no hybridisation for* ***S15-D$_1$ /T0****, EtBr does not bind and does not show up in the image.*

3.6 Circular Dichroism

3.6.1 Introduction to Circular Dichroism

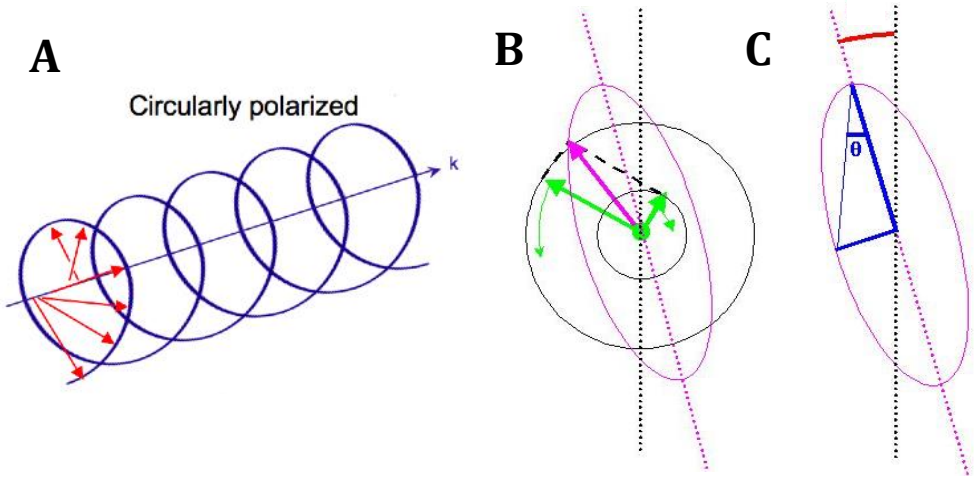


Figure 18. (A) A plot of a circularly polarised wave, with the electric field vector (red) drawing out a helix (blue) as the wave propagates. (B) Upon the differential absorption of circularly polarised light an elliptical wave is produced, where the overall vector (magenta) is a sum of the now unequal vectors of left and right circularly polarised light (green). (C) A spectrapolarimeter measures the ellipticity (θ) of this wave as a function of wavelength.

Circularly polarised light is an optical property whereby the electric field vector of an electromagnetic wave does not change in magnitude but rather changes direction in a circular motion as the wave propagates. The direction of the electric field vector can travel in either a clockwise (right handed) direction, or in an anti-clockwise (left handed) direction, and will plot a helix along the direction of propagation (Figure 18A).[18]

Circular dichroism (CD) is a technique that measures the differential absorption of left and right circularly polarised light (ΔA_L-A_R) as it travels through an optically active medium, since circularly polarised light itself is said to be chiral, the two forms interact with chiral compounds differently. A spectrapolarimeter passes both right and left circularly polarised light through a sample simultaneously, if the sample has a different absorbance for the right polarised light in comparison

to the left the resulting wave becomes elliptically polarised (Figure 18B-C). A CD spectrum displays the ellipticity (θ) of the elliptical wave as a function of wavelength (λ).[19]

3.6.2 Applications of Circular Dichroism

Circular dichroism spectroscopy has found many uses in a variety of fields, but most notably it is possible to utilise UV-CD to investigate the secondary structure of proteins and nucleic acids. Circular dichroism of DNA arises from the asymmetric sugars within the backbone as well as the helical structure of the duplex, which gives DNA its chirality. In the accessible region of the UV-CD spectrum (200 – 300 nm), the absorbance is dominated by π-π* transitions of the nucleobase, which if measured independently exhibit no signal as they are achiral. Therefore the measured CD transitions of the nucleobases are a result of their coupling with the backbone transitions. There is also a contribution from stacking interactions within polynucleotides, with increased magnitude at 270 nm due to the bases stacking in a chiral (helical) fashion.[20]

Each polymorph of DNA (A, B & Z) has its own characteristic set of CD signals (Figure 19), since each of them has their own unique stacking patterns. This allows them to be distinguished from each other through CD spectroscopy and the technique has been widely used to monitor the transitions between each type in response to a number of factors such as pH or solvent. The B-DNA type structure, which is the most common and is the structure adopted by the anthracene

modified oligonucleotide, has two characteristic bands; one negative and one positive at 240 and 270 nm respectively.[21]

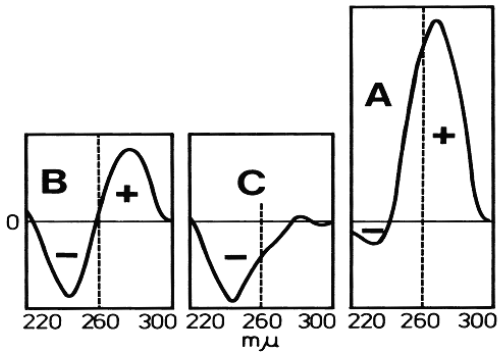


Figure 19. Comparison of CD bands present when analysing the various DNA polymorphs (A, B and C/Z).[20]

Molecules that bind and intercalate within the DNA structure will also have an induced CD signal even if they were not chiral structures to begin with, making circular dichroism an ideal technique for measuring binding constants, as well as giving an indication as to whether the binding leads to any disruption of the DNA secondary structure.[22] The induced CD effect was also observed for anthracene tagged sequences in previous research, where observations from CD spectroscopy were used to correlate with data from other techniques to confirm the position of the tagged anthracene with respect to the overall DNA structure. In comparison to unmodified DNA, the anthracene tag introduces a shoulder peak within the CD spectrum at 254 nm due to an induced chirality from being covalently bound to the DNA backbone. This is in combination with stacking interactions with neighbouring bases, but only if the anthracene is capable of adopting an intercalated position.[23]

When analysed through CD spectroscopy, sequences **S14**-**S16** each displayed their own characteristic set of CD bands in comparison to the unmodified

sequence **S0** as well as compared to each other, but all exhibited the positive and negative bands indicative of the B-DNA type structure, (Figure 20). The introduction of the anthracene modification appears to have disrupted the overall stacking interactions within the duplex, which results in a reduction in magnitude of the positive band at 270 nm compared with the unmodified sequence. This can be explained by the fact that when an anthracene intercalates it can force the DNA to unwind and since the band at 270 nm is an accumulation of all base pair interactions there is an observed decrease.[23] The induced anthracene band at 254 nm, which appears as a shoulder on the positive band, differs from sample to sample with varying intensity, but this is most likely not due to the proximity of each anthracene tag with respect to each other but due to which neighbouring bases flank the tag within each sequence. The ability of anthracene to intercalate and therefore show up in the CD spectrum depends upon the π-stacking interactions to a neighbouring base, therefore the magnitude of the anthracene

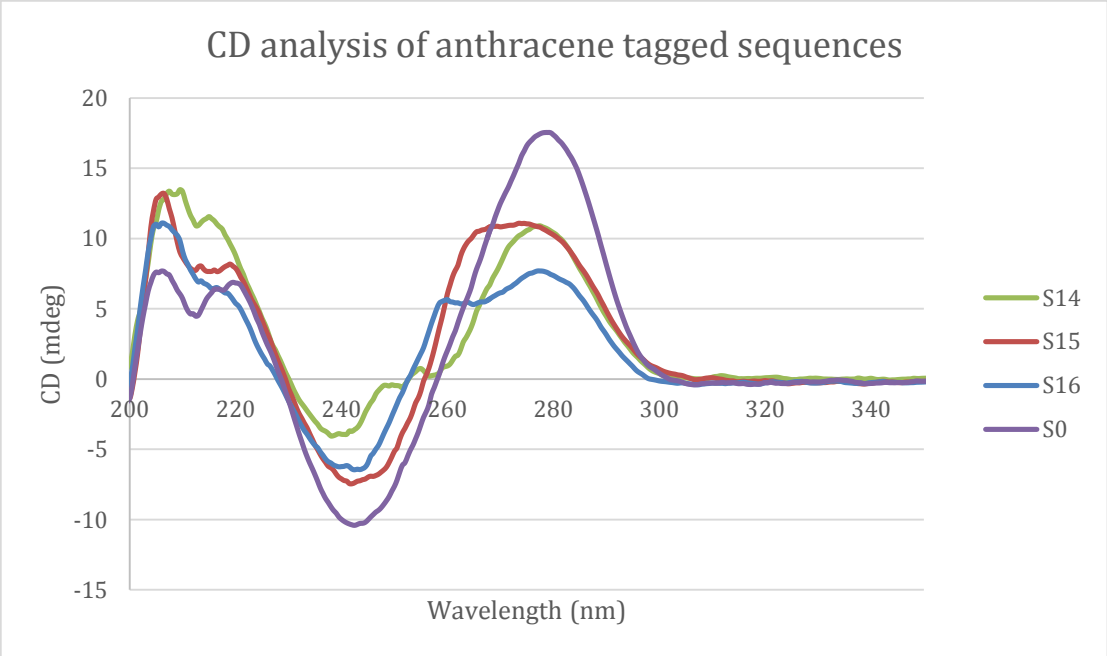


*Figure 20. CD analysis of anthracene modified oligonucleotides **S14**, **S15** and **S16** as duplexes formed with the unmodified sequence **S0**. The anthracene modification appears as a shoulder at 250 nm.*

band in sequences **S14**-**S16** is a reflection of the differing interactions that occur when the neighbouring bases change from sample to sample.

3.6.3 Circular Dichroism and Photo-Switched Binding

Single stranded DNA is structurally less well defined than duplex DNA and so their CD signal is smaller. CD spectroscopy can therefore be used to differentiate between SS and DS states and can confirm earlier observations made through thermal melting point analysis. Each sequence was analysed through CD spectroscopy before and after dimerisation in order to observe what information can be gathered about the secondary structure of the anthracene modified oligonucleotide upon photodimerisation.

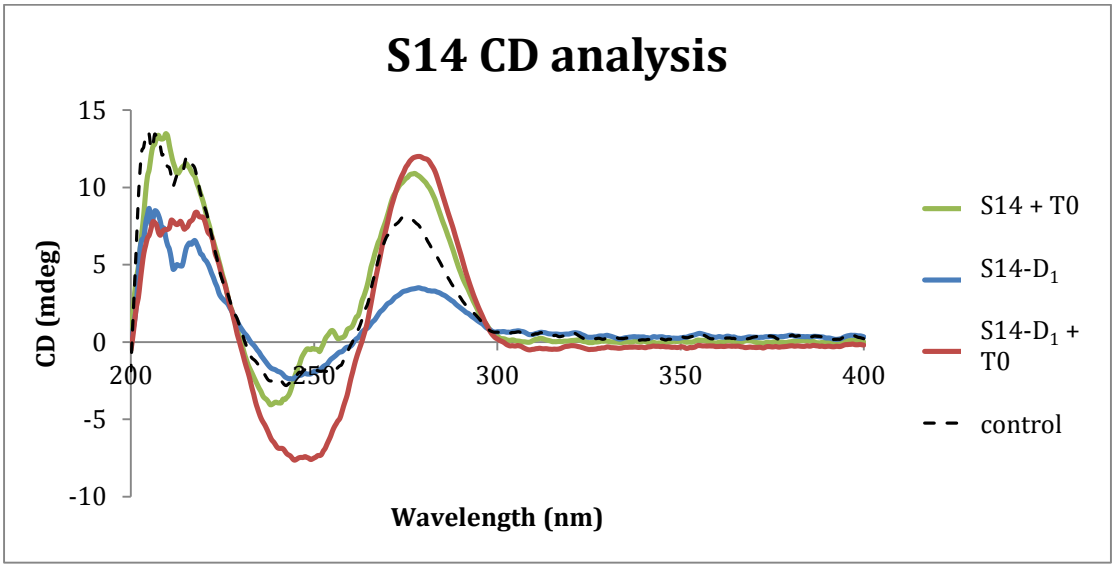


*Figure 21. CD analysis of sequence **S14**, there is no significant change in duplex stability upon photodimerisation, therefore only a change in the induced CD band of anthracene is observed. Control and experimental data do not match, which also confirm this observation.*

As anticipated the CD bands for sequence **S14** remain largely unchanged after photodimerisation as there was no significant change in duplex stability and the B-DNA bands are still present. The only change in the CD after dimerisation is that

the induced anthracene band at 250 nm disappears, which is to be expected as the photodimer is no longer able to take part in π-stacking interactions which give rise to this band (Figure 21).

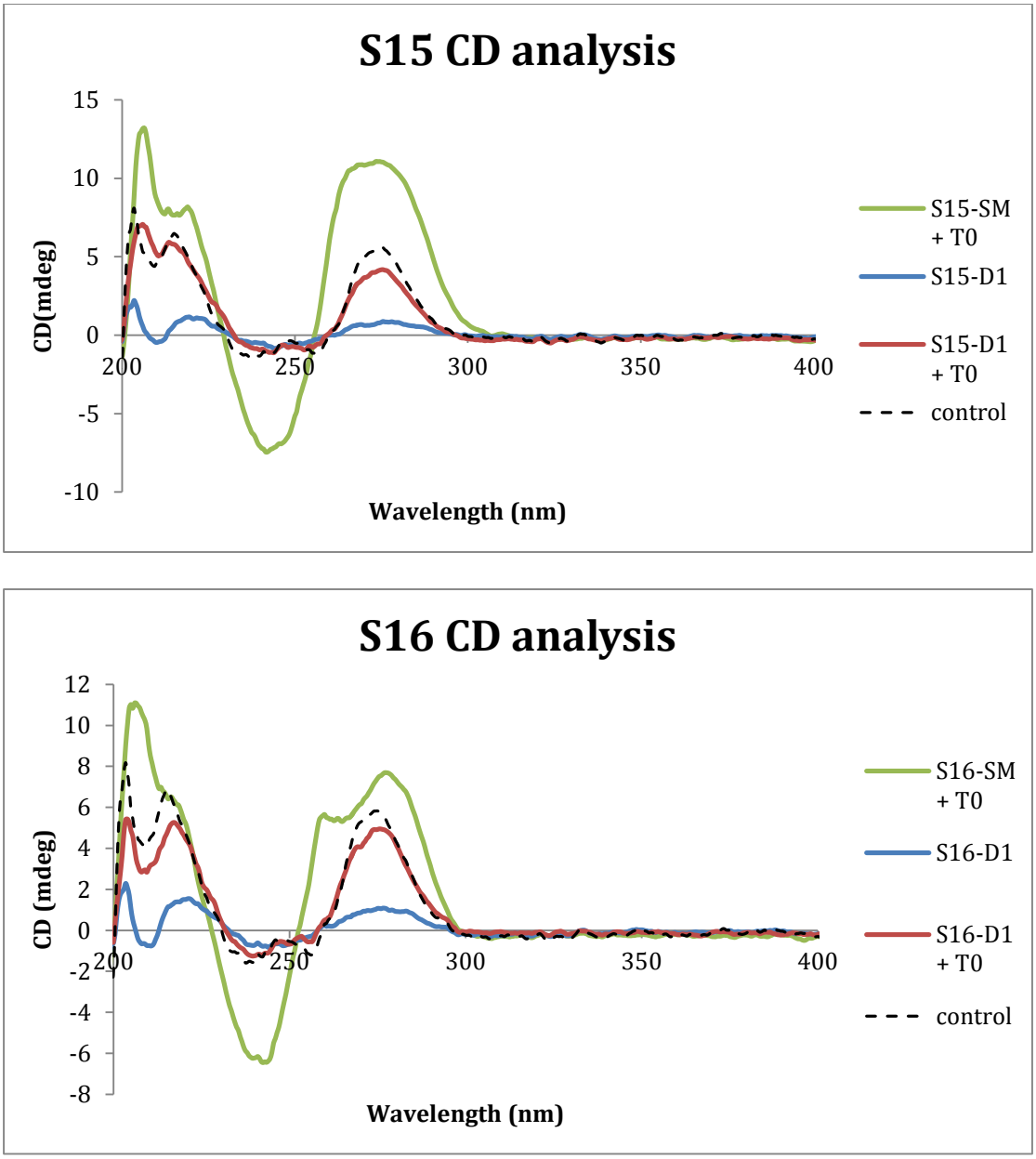


*Figure 22. CD analysis of sequences **S15** and **S16**, significant changes in the CD bands were observed upon photodimerisation. Control CD data collected from scanning the strands independently confirm no hybridisation has occurred.*

As for sequences **S15** and **S16** which do show a drop in thermal stability after photodimerisation there was a significant change in the CD spectrum observed. There was a reduction in both B-DNA bands, in particular the negative band at 245

nm which correlates to duplex helicity, suggesting that the double helix has not been formed (Figure 22). As a control measure the D_1 photodimer and the complementary strand (**T0**) were scanned individually for each sample and the data was mathematically added together to give a new control trace (displayed as a black dashed line), which has been overlaid onto each spectrum. The control trace is used to show how the CD spectrum should ideally look if each strand were in a single stranded state with zero interaction, and for sequences **S15** and **S16** these data are comparable with each other, confirming the results. As for sequence **S14** the data sets do not match, again confirming that in this case photodimerisation does not inhibit duplex formation and an interaction is still possible.

3.7 Reversion Studies

In order to reversibly control hybridisation using anthracene photodimerisation, the process of dissociating the photodimer to reform the starting material needed to be fully investigated. Dissociation of an anthracene dimer is possible in two ways; application of heat or through irradiation with UV light of a shorter wavelength than dimerisation (250 to 290 nm).[24] The photodissociation of anthracene is known to involve a photostationary state and also requires UV light in the same absorbance range as DNA (260 nm), meaning irradiation at this wavelength would be unfavourable as it could initiate other photochemical reactions.[2,25] Therefore the primary method of reversion used in this study has been through thermal dissociation, which was analysed by measuring the absorbance at a range of temperatures using a VT-UV spectrometer to detect the re-emergence of the anthracene band (360 nm) as the dimer breaks down.

On first analysis, it was found that the D_1 photodimers of samples **S14**-**S16** were reasonably stable to heat, with very minor changes in the absorbance at temperatures ranging from 30 to 50 °C. Only when the temperature was elevated to 80 °C or higher was there an observable change and even under this degree of heating the full reversion of the photodimer progressed over a period of 12 hours. Complete reversion of the photodimer for each sequence was confirmed through analytical HPLC, where the D_1 peak was seen to gradually disappear, replaced by the starting material peak at the longer retention time.

In order to quantify the rate of dissociation (k_{diss}) the absorbance was measured at set intervals during reversion until the endpoint was reached. The dissociation constants were calculated using the following relationship in equation 1.[26]

Equation 1. $$k_{diss} = -\frac{1}{t}\ln\left[\frac{(A_\infty - A_t)}{(A_\infty - A_0)}\right]$$

t = Time
A_t = Absorbance at time t during reversion
A_∞ = Absorbance at reversion endpoint
A_0 = Absorbance prior to reversion
k_{diss} = rate of reversion

Plots of $\ln[(A_\infty - A_t) / (A_\infty - A_0)]$ against time gave a straight line (Figure 23), where the slope provided the thermal dissociation rate constant (k_{diss}). The dissociation constant for sequences **S14-D$_1$**, **S15-D$_1$** and **S16-D$_1$** were determined to be 2×10^{-3}, 1.9×10^{-3} and 2.6×10^{-3} s^{-1} respectively, which indicates that there may be some relationship to the base separation between the anthracene tags. Whilst sequences **S14-D$_1$** and **S15-D$_1$** display comparable dissociation rates, by extending the number of bases between tags to five in sequence **S16** the rate of

dissociation is increased, suggesting that in this configuration the anthracene photodimer is placed under more structural strain.

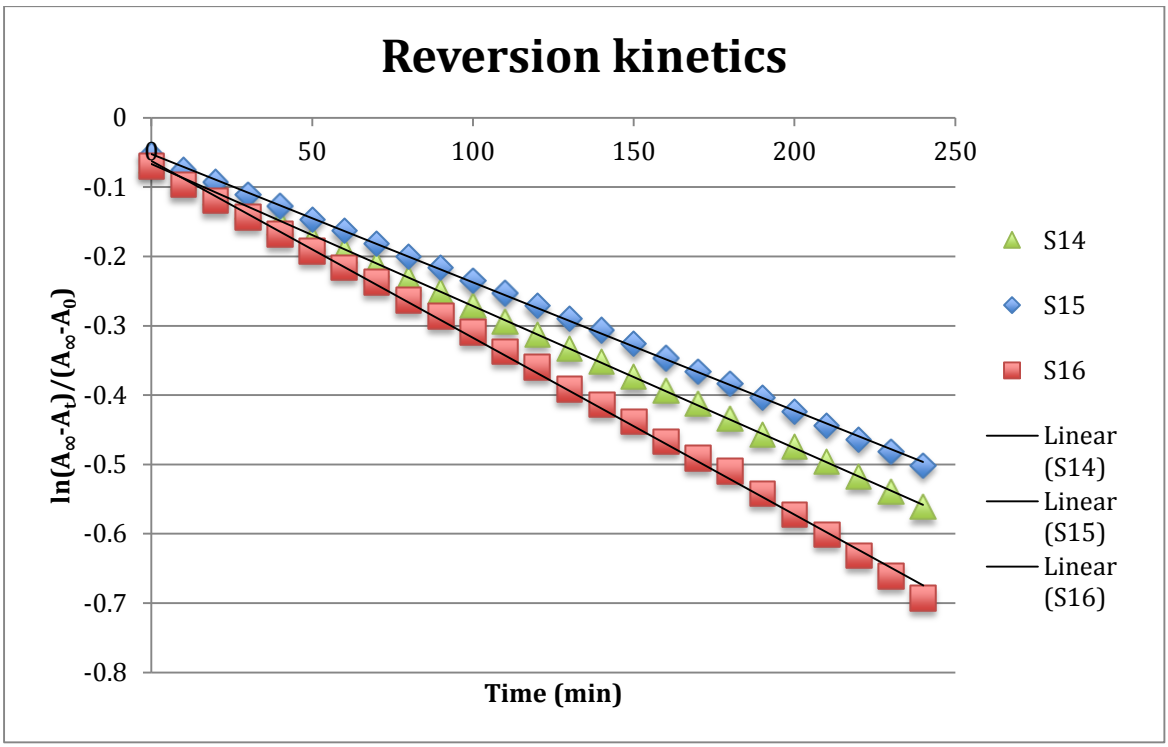


Figure 23. A plot of thermal reversion data used to determine the dissociation constant (k_{diss}) using the relationship described in equation 1. Samples heated to a constant temperature of 80 °C recording at 1 minute intervals at 260 nm.

Due to the prolonged period over which the D_1 photodimer reverted to the starting material, its progress could be monitored through T_m analysis. Figure 24 displays T_m curves recorded at different stages of heating the sample, and as reversion progresses the curve shifts towards a higher temperature. At intermediate stages T_m transition curves can be seen for both the dimer and starting material.

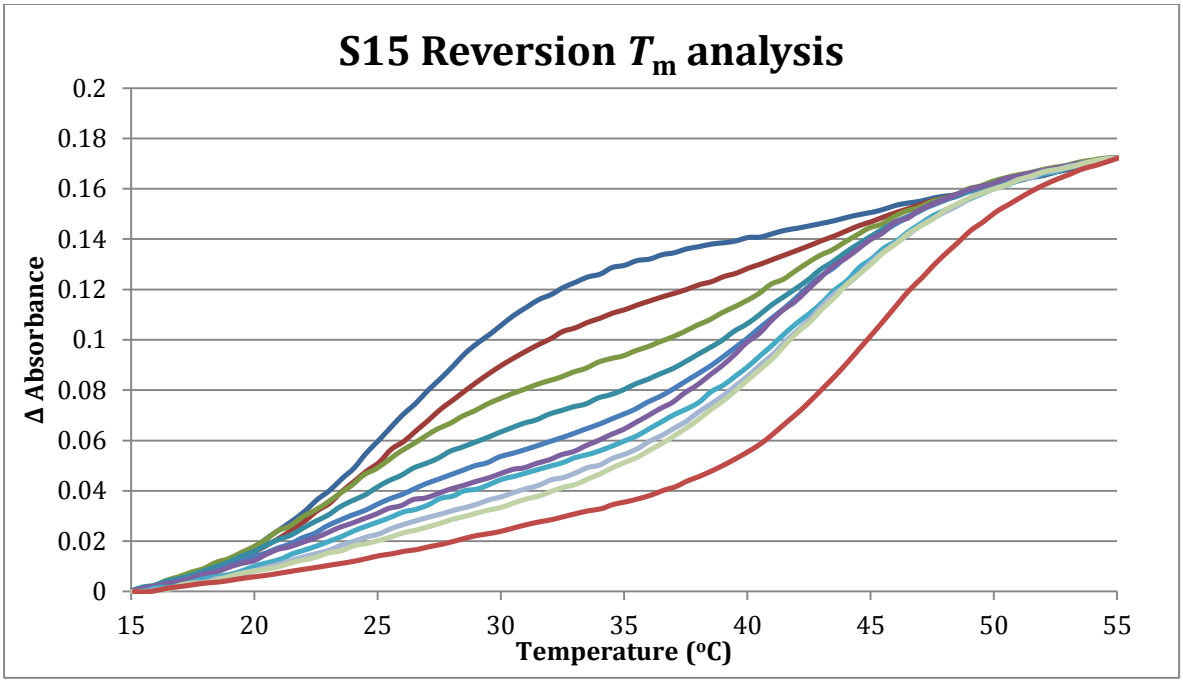


Figure 24. Ten T_m ramps collected progressively, which displays the effect of the thermal reversion of the anthracene photodimer. The T_m transition curve was seen to shift gradually from that of the D_1 photodimer to that of the original starting material.

3.8 In Situ Irradiation and Gated Photochromism

An important aspect of this study is being able to achieve what has been termed 'photo-release', where the anthracene modified sequence is irradiated 'in-situ' in the presence of the complementary strand. The initiation of the photodimerisation reaction would then drive the sequence to switch between double and single stranded states, providing reversible control over hybridisation and in effect releasing the bound complementary strand with the use of light.

In order to test the theory of 'photo-release', sequences **S14**, **S15** and **S16** were mixed with an equal ratio of the complementary strand **T0** in a buffered solution and annealed prior to irradiation in order to form the double stranded duplex. Figure 25 displays the reaction profiles of the double stranded samples, monitored at 365 nm over a period of 40 minutes. The data indicated that whilst

sample **S14** did show signs of photodimerisation, the reactivity of samples **S15** and **S16** was significantly reduced compared to those in the single stranded state (Figure 25). Sample **S16** displayed little to no change over the course of the reaction and appeared to be completely suppressed in the double stranded state. From earlier CD research it is known that the anthracene tags within the modified oligonucleotide have a preference to intercalate between neighbouring bases due to their planar aromatic structure. This preference could potentially render the anthracene tags immobile when paired with the complementary sequence, precluding their availability for photodimerisation. Another important factor to note is the relative positions the anthracene tags occupy once the B-DNA helix structure is formed. There are 10.5 base pairs per helical turn of DNA, therefore as the number of bases between each anthracene tag is increased then the angle of orientation of each tag with respect to each other is also increased, (Figure 26). For sequence **S14** the degree of rotation between the tags remains relatively small, however for sequence **S16**, with a 5 base pair separation, the anthracene tags end up on opposite sides of the double helix structure (206°). Couple this with the vertical distance separating the tags (3.4 Å per base pair) and it is clear that the formation of a double helix can severely limit the proximity between anthracene tags and therefore supress photodimerisation.

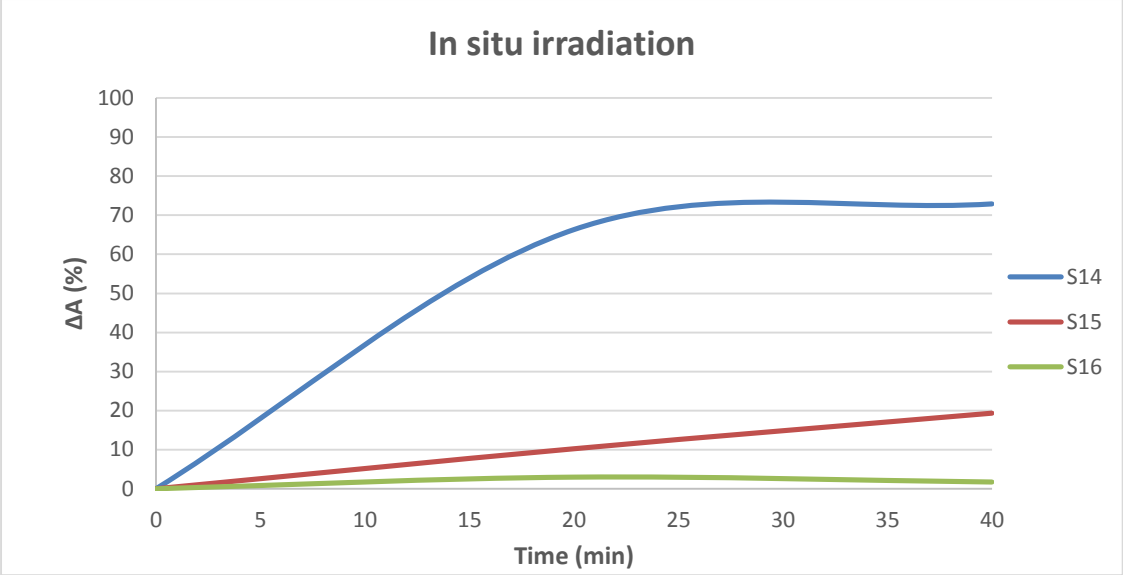


*Figure 25. In situ irradiation of sequences **S14**, **S15** & **S16** in the presence of an excess of the complementary strand **T0**.*

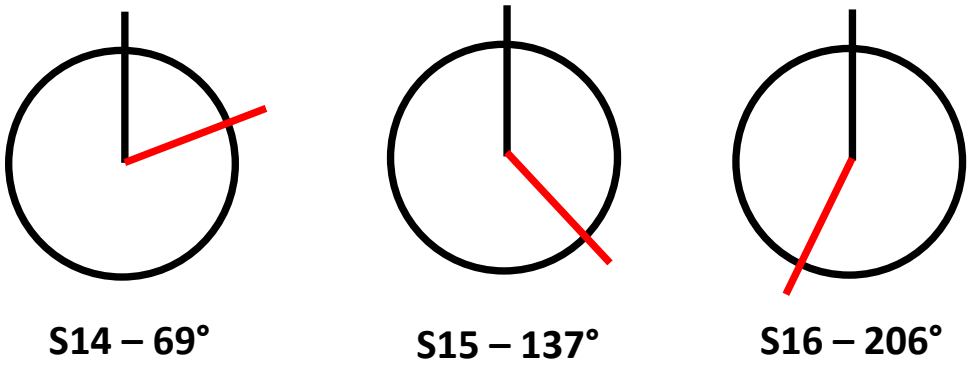


Figure 26. Angles of rotation about the vertical axis of the B-DNA double helix between anthracene tags. As the number of base pairs between anthracene tags increases so does the angle rotation and therefore distance between groups within the duplex structure.

Whilst samples **S14** and **S15** did show signs of reactivity in the presence of the complementary strand the T_m data collected for these sequences suggest that there would still be sufficient hybridisation post irradiation. Therefore, even if sufficient yields of the photodimer could be obtained in these in-situ experiments, it may not necessarily mean that the complementary strand is 'released' through photodimerisation.

The interesting aspect of the in-situ reactivity of sequence **S16** is that it displays what has been termed 'gated photochromism'. The presence of the complementary strand clearly inhibits photodimerisation by limiting the proximity of the anthracene tags, and from earlier T_m data it was known that the successful dimerisation of **S16** subsequently blocks hybridisation to the complementary **T0**. Therefore much like an electrical 'gated' switch the system is sensitive to two inputs, in this case one optical and one chemical. Each input is capable of switching off the process associated with the other and in essence the

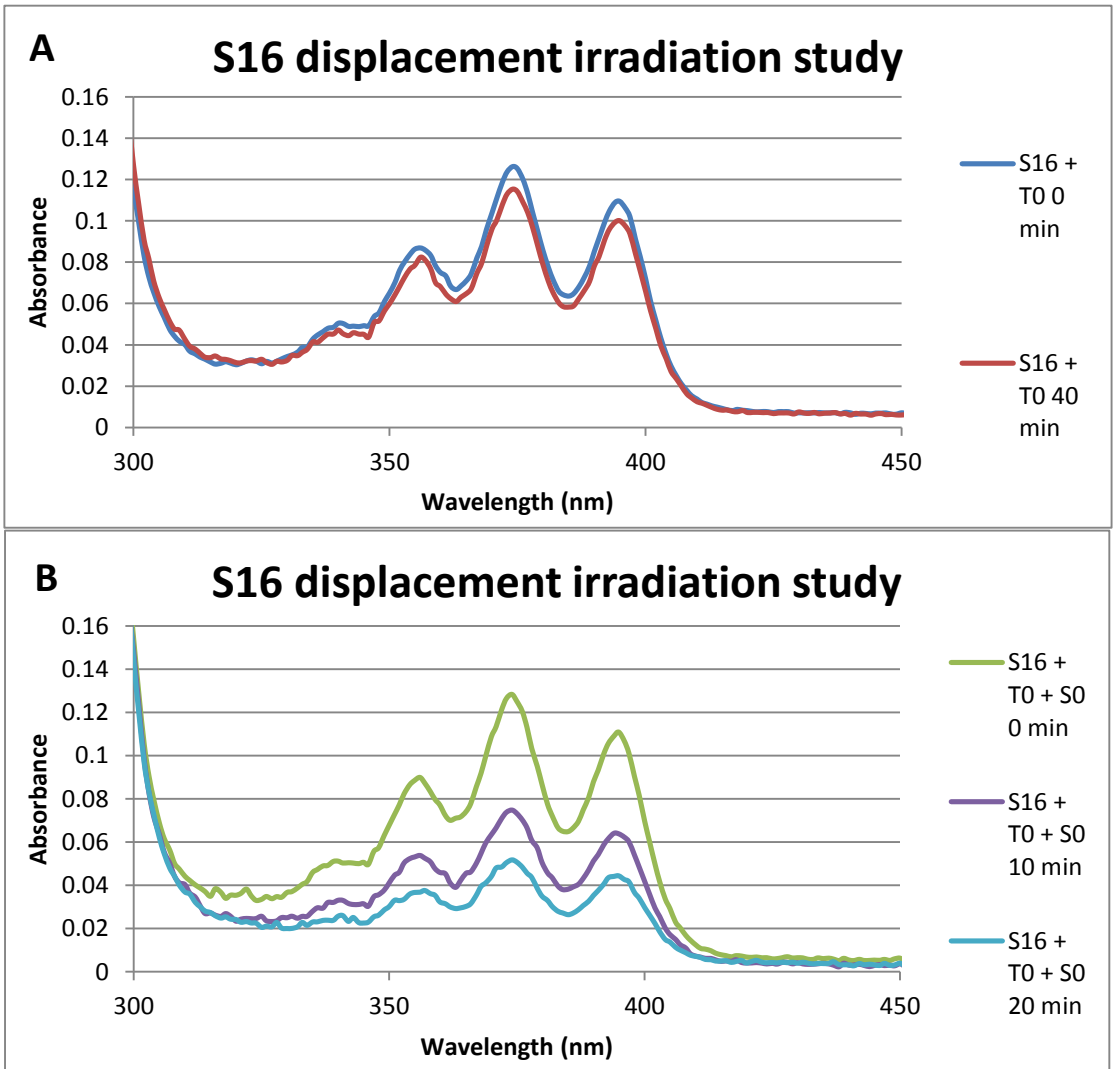
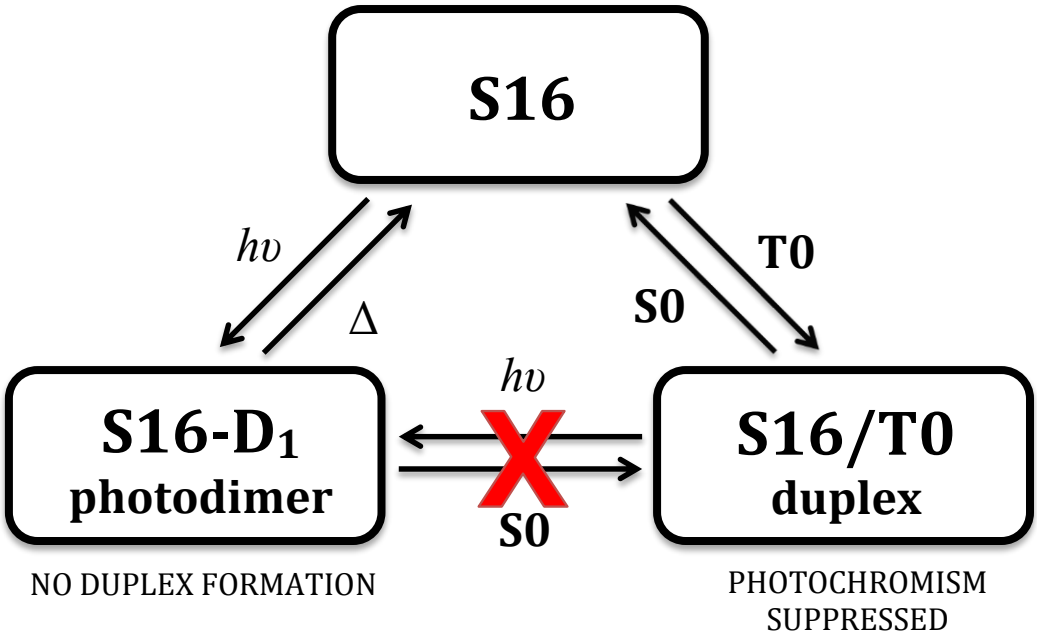


*Figure 27. Irradiation studies indicating that (A) photodimerisation is suppressed by the presence of the complementary strand **T0**, with little change in absorbance during irradiation. (B) Only after sequence displacement using the unmodified sequence **S0** can dimerisation resume.*

system displays both photocontrolled duplex formation and binding-controlled photochromism, (Scheme 5). It was found however that although the presence of **T0** blocks dimerisation (Figure 27A), when the unmodified sequence **S0** was added it is capable of competing with the anthracene modified strand to form the more stable **S0**/**T0** duplex through displacement. This in effect releases the **S16** sequence and upon a secondary irradiation trial, the photodimerisation reaction was restored, (Figure 27B). Therefore it is possible to move between the duplex and photodimer forms of the oligonucleotide through the two reversible processes as shown in Scheme 5.

The repercussions of this observation severely limit the applications in which anthracene photodimerisation can be used to control hybridisation and significant efforts are required to be adapt this method in order to overcome the issues discussed in this chapter.

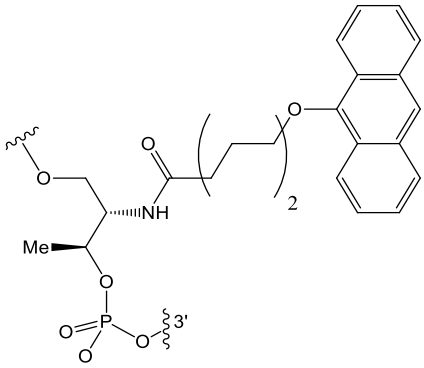


Scheme 5. Schematic representation of gated photochromism. It is possible to shift indirectly between the single stranded and duplex states using a combination of processes, however it is not yet possible to initiate photo-release using anthracene photodimerisation alone.

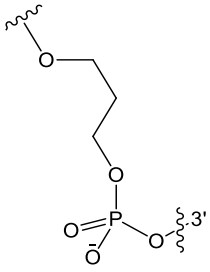
3.9 Control Experiments

As seen in earlier work by Asanuma and Komiyama, the removal of a nucleobase from a DNA sequence can significantly reduce the overall stability of the duplex structure due to the reduction in π-stacking and hydrogen bonding interactions. However replacing the nucleobase with a synthetic group that is both planar and aromatic can compensate for this change, and in the case of azobenzene in Asanuma's research it can over-compensate, as the overall stability was seen to increase.[27] This is due to the favourable π-stacking interactions that occur between the azobenzene and the neighbouring bases when it is in an intercalated position. Anthracene also has a large, planar aromatic structure and is known from earlier research by Duprey to favour an intercalated position, depending on the type of linker used to incorporate the group into the oligonucleotide. Therefore it is important to verify the changes that are induced when replacing two bases within an oligonucleotide with anthracene tags, and how intercalation compensates for the change in the supramolecular interactions.

When the two anthracene tags within an oligonucleotide are irradiated and are converted to the photodimer, they will no longer be able to adopt an intercalated position as the photodimer has a non-planar structure. Therefore the reduction in duplex stability upon photodimerisation can be partially attributed to the loss in π-stacking between anthracene and its neighbouring bases. However as seen from earlier T_m analysis, there is a relationship between the base separation and photodimer duplex stability that will be due to a structural alteration. Consequently it is also important to investigate these competing factors and how each of them contribute to the overall reduction in the T_m.



Anthracene monomer structure (**X**)



Propyl (C3) spacer (**Y**)

Figure 28. Structure of the anthracene monomer and the control C3 spacer group used to replicate the structure with the exclusion of the appended anthracene.

To explore these important issues, a series of control experiments were developed where the anthracene within each sequence (**S14-16**) is systematically replaced with a propyl (C3) spacer, which has a structure that is analogous to the anthracene monomer, with the exception of the anthracene attachment (Figure 28).[28] Table 6 outlines the control sequences used in this experiment, where **X** denotes an anthracene tag and **Y** a propyl spacer. Note that three controls have been created for each sequence, with one abasic replacement in **A** & **B** and two replacements in **C** in order to fully investigate how anthracene intercalation contributes towards the stability of the duplex. By comparing the T_m data from these control sequences with earlier results, as well as that of the unmodified sequence, the contributions from π-stacking interactions and structural alterations can be better understood.

*Table 6. Sequences of the nine control oligonucleotides incorporating the C3 control spacer (**Y**) and anthracene monomer (**X**).*

Control	S14	S15	S16
A	TGGACT**X**T**Y**TCAATG	TGGAC**X**CTC**Y**CAATG	TGGA**X**TCTCT**Y**AATG
B	TGGACT**Y**T**X**TCAATG	TGGAC**Y**CTC**X**CAATG	TGGA**Y**TCTCT**X**AATG
C	TGGACT**Y**T**Y**TCAATG	TGGAC**Y**CTC**Y**CAATG	TGGA**Y**TCTCT**Y**AATG

*Table 7. T_m data of control sequences **A**-**C** compared with previous data from starting material (**SM**) and dimer (**D1**) samples. Data from control sequences gives an indication of how anthracene intercalation contributes to duplex stability. Calculating the value Δ(T_m**C** - T_m**D1**) determines to what extent structural changes alter the T_m.*

	S14	S15	S16
SM	37.5	45	35
D1	37	25	< 5
A	38	40	28.5
B	37.5	40	27
C	30	31.5	19.5
Δ(T_mC-T_mD1)	-	6.5	> 14.5

When all the T_m data from the control sequences had been collected and cross analysed there was a clear trend observed as each anthracene was systematically replaced by the propyl spacer. With each subsequent replacement there was a drop in T_m representative of the loss in π-stacking interactions when the anthracene is removed. If we consider the data for **S15**, comparing the values for the starting material **S15** and control sequence **S15-C,** there is a 13.5 °C reduction which can be attributed to the loss in π-stacking interactions. When compared with the T_m value of the photodimer **S15-D1**, this accounts for approximately two thirds of the overall change in duplex stability upon photodimerisation, with the structural alterations therefore responsible for the remaining one third drop in the T_m.

One exception to this trend is sequence **S14**, where there was no significant reduction in the T_m upon photodimerisation. What was unexpected was that when both anthracene tags within this sequence were substituted with the propyl spacer, there was a 7.5 °C drop in the T_m, compared with a 0.5 °C drop upon photodimerisation. This would seem to suggest that even in the form of the photodimer, there is still some additional stabilisation, of which can only be explained by the proximity of each tag within this sequence. It is feasible that the anthracene photodimer can still adopt a position within the hydrophobic core of the duplex structure if the base sandwiched between the two tags is ejected, leaving a large enough space to accommodate the photodimer. In this case the insertion of the photodimer within the duplex could lead to a more stable structure and therefore a higher T_m than its double abasic counterpart **S14-C**.

3.10 Fluorescent Analysis

A secondary observation for sequence **S14** was that there is no significant change in T_m between **S14** and **S14-A/B**, which would suggest that they potentially have similar structures, having only one intercalated anthracene in each case. This could again be possibly explained by the relative proximity of the two anthracene tags within this sequence. When anthracene is intercalated, neighbouring bases can be forced off axis in order to accommodate it as seen in the schematic drawing of Figure 29. In this case the intercalation of one anthracene restricts the cavity within the duplex in which the second tag would have been located, therefore preventing intercalation.

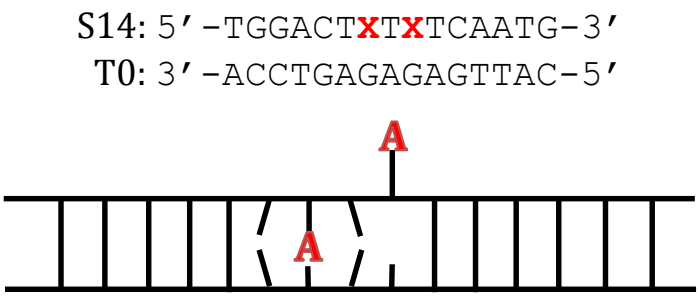


*Figure 29. A schematic illustration of the anthracene (**A**) intercalation in the **S14/T0** duplex. Distortion of neighbouring bases prevent intercalation of the secondary anthracene due to their proximity in this sequence. This effect may cause differences in the T_m and fluorescence data from what was expected due to the anthracene not being intercalated.*

In order to determine that this was the case a series of experiments were carried out to monitor the fluorescence of anthracene as the sequence was titrated against the complementary sequence. Anthracene is known to have an increased fluorescence signal when intercalated between base pairs within double stranded DNA due to a reduction in solvent quenching. Therefore comparing the increase in fluorescence between samples would give an indication as to whether both anthracene tags have successfully intercalated. Figure 30 displays a normalised fluorescence spectrum of sequences **S14** and **S15**; when titrated against the complementary sequence **T0** there is a corresponding increase observed in the fluorescence signal. Comparing the fluorescence between samples following the addition of one equivalent of the complementary strand it is clear that the increase in emission for sequence **S14** is less than half that of **S15**. This would seem to confirm that the anthracene tags within sequence **S14** are not capable of taking part in successive π-stacking interactions as well as other samples, explaining the earlier observation in T_m data.

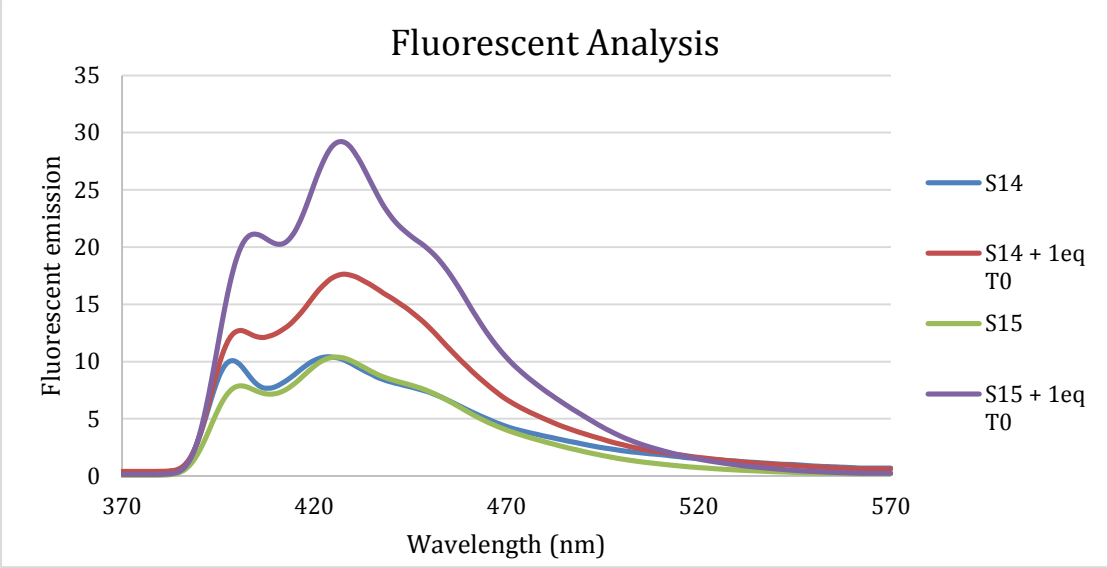


*Figure 30. Fluorescent analysis of sequences **S14** & **S15** when titrated against **T0**. The decreased fluorescence of **S14** compared with **S15** when titrated with **T0** would suggest only one of the anthracene modifications is capable of intercalating, whilst the other remains quenched by surrounding solvent.*

S14	***S15***
TGGACT**X**T**X**TCAATG	TGGAC**X**CTC**X**CAATG
ACCTGAGAGAGTTAC	ACCTGAGAGAGTTAC

3.11 Conclusions and Further Work

Within this chapter it has been demonstrated that the hybridisation interaction between complementary sequences of DNA can be reversibly controlled with the use of the bimolecular photodimerisation reaction. Whilst the technique has displayed some distinct advantages such as the ON/OFF binding behaviour of sequence **S14** and the minimal disruption that is involved when inserting the synthetic anthracene groups, this research has thrown up some challenges, which still need to be addressed. For example the rate of dissociation is too slow for use in practical applications and requires excessive heating to achieve full reversion, and while 'gated' photochromism is an interesting property, the ability to achieve photo-release of complementary sequences remains a desirable target for this research.

Several strategies that have been conceived in an attempt to overcome the intercalation of anthracene that leads to gated photochromism are discussed in subsequent chapters. However, a number of proposals that require the synthesis of new anthracene monomers could also make photo-release possible. For example, if the anthracene group is adapted with additional functional groups such as a methyl or tertiary butyl group on the C10 position, then it would no longer be planar and would potentially disfavour intercalation. In addition to this, the bulky alkyl groups would put the head to tail photodimer under more structural strain and therefore may improve dissociation kinetics during thermal reversion (Figure 31).

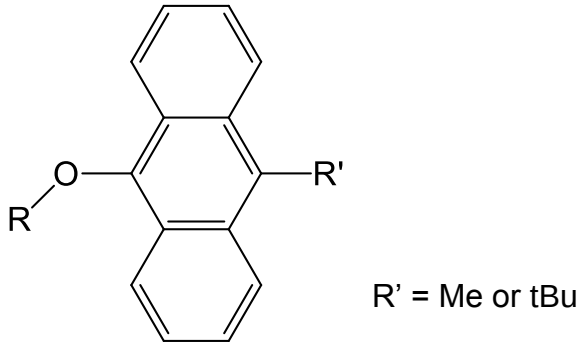


Figure 31. Structure of a new anthracene phosphoramidite, where R' on the C10 position introduces a bulky group that could prevent DNA intercalation.

Alternatively the threoninol linker that connects the anthracene to the DNA backbone could be replaced. Following on from other research within the field by Maeda and Jäschke where photochromic groups are attached or incorporated into the structure of a nucleobase (discussed in Section 1.7.2 and 1.7.3 respectively) the anthracene phosphoramidite could be synthesised to include a nucleobase within its design.[29,30] The nucleobase would potentially take part in both hydrogen bonding and π-stacking interactions to the adjacent base on the complementary

strand, filling the gap where the anthracene would normally intercalate. This strategy would in theory improve results as the introduction of an anthracene modification would not result in a net reduction in hydrogen bonding as it does with a threoninol linker. This type of strategy has already been attempted by the Saito group, which were able to modify a uridine nucleotide with anthracene through a Sonogashira coupling for use in base-discriminating fluorescent oligonucleotide probes (Figure 32).[31] One of the major advantages of the nucleotide developed by Saito is that the linker contains a rigid alkyne group, which could potentially help orient the anthracene away from the duplex. Therefore if the nucleotide was adapted to include additional carbons within the linker it could be used to prevent anthracene intercalation once incorporated into an oligonucleotide and potentially work towards a method of anthracene mediated photo-release. The major downside to this strategy is that multiple phosphoramidites would have to be synthesised to include different nucleobases (A, G, C & T) in order to be able to synthesise sequences with differing sequence composition.

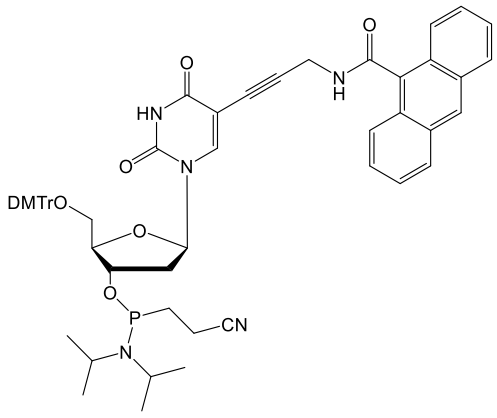


Figure 2. An anthracene modified uridine phosphoramidite developed by Saito et al. for the purpose of base discriminating fluorescent oligonucleotide probes

3.12 References

1. Jabłoński, A., *Nature*, **1933**, *131*, 839-840.
2. Bouas-Laurent, H.; Castellan, A.; Desvergne, J.; Lapouyade, R., *Chem. Soc. Rev.*, **2001**, *30*, 248-263.
3. Woodward, R. B.; Hoffman, R., *Acc. Chem. Res.*, **1968**, *1*, 17-22.
4. Fleming, I., *Oxford Chemistry Primers: Pericyclic Reactions*., Oxford University Press, Oxford, **1998**, Chapter 3, 31-56.
5. Applequist, D. E.; Friedrich, E. C.; Rogers, M. T., *J. Am. Chem. Soc.*, **1959**, *81*, 457-458.
6. Sigman, M. E.; Zingg, S. P.; Pagni, R. M.; Burns, J. H., *Tett. Lett.*, 32, 5737-5740.
7. Duprey, J.-L. H. A.; Bassani, D. M.; Hyde, E. I.; Ludwig, C.; Rodger, A.; Vyle, J. S.; Wilkie, J.; Zhao, Z. Y.; Tucker, J. H. R., *Supramol. Chem.*, **2011**, *23*, 273-277.
8. Duprey, J.-L. H. A., (**2010**), *Studies on Anthracene Tagged Oligonucleotides*, Ph.D. thesis, University of Birmingham, UK.
9. Aldaye, F. A.; Palmer,A. L.; Sleiman, H. F., *Science*, **2008**, *321*, 1795-1799.
10. Jones, P., *Gel Electrophoresis: Nucleic Acids*, Wiley Publishing, Chichester **1995**.
11. Rickwood, D.; Hames B. D., *Gel Electrophoresis of Nucleic Acids: A Practical Approach*., IRL Press Ltd, **1982**.
12. Sanger, F.; Coulson, A. R., *J. Mol. Biol.*, **1975**, *94*, 441-448.
13. Frank, R.; Müller, D.; Wolff, C., *Nucleic Acids Res.*, **1981**, *9*, 4967-4979.
14. Zhao, Z. Y.; Wilson, T. J.; Maxwell, K.; Lilley, D. M., *RNA*, **2000**, *6*, 1833-1846.
15. Hassur, S. M., Whitlock, H. W., *Anal. Biochem.*, **1974**, *59*, 162-164.
16. Waring, M. J., *J. Mol. Biol.*, **1965**, *13*, 269-282.
17. Higuchi, R.; Dollinger, G.; Walsh, P. S.; Griffith, R., *Nat. Biotechnol.*, **1992**, *10*, 413-417.
18. Kliger, D. S.; Lewis, J. W.; Randall, C. E., *Polarized Light in Optics and Spectroscopy*, Elsevier, **1990**.
19. Rodger, A.; Nordén, B., *Circular Dichroism & Linear Dichroism*, Oxford Chemistry Masters, **1997**.
20. Gratzer, W. B., *Eur. J. Biochem.*, **1970**, *15*, 209-214.
21. Kypr, J.; Kejnovská, Renčiuk Vorlíčková, M., *Nucleic Acids Res*., **2009**, *37*, 1713-1725.
22. Nordén, B.; Kurucsev, T.; *J. Mol. Recognit.*, **1994**, *7*, 141-156.
23. Rodger, A.; Taylor, S.; Adlam, G.; Blagbrough, I. S.; Haworth, I. S., *Bioorg. Med. Chem.*, **1995**, *3*, 861-872.
24. Grimme, S.; Peyerimhoff, S. D.; Bouas-Laurent, H.; Desvergne, J.-P.; Becker, H.-D.; Sarge, S. M.; Dreeskamp, H., *Phys. Chem. Chem. Phys.*, **1999**, *1*, 2457-2462.
25. Sinha, R. P.; Häder, D.-P., *Photochem. Photobiol. Sci.*, **2002**, *1*, 225-236.

26. McSkimming, G., *Photo-active Anthracene Receptors for s-Block and d-Block Metals,* University of Exeter, **2001**.
27. Kashida, H.; Liang, X.; Asanuma, H., *Curr. Org. Chem.*, **2009**, *13*, 1065-1084.
28. Takeshita, M.; Chang, C.-N.; Johnson, F.; Will, S.; Grollman, A. P., *J. Biol. Chem.*, **1987**, *262*, 10171-10179.
29. Ogasawara, S.; Maeda, M., *Angew. Chem. Int. Ed.*, **2008**, *47*, 8839-8842.
30. Singer, M.; Jäschke, A., *J. Am. Chem. Soc.*, **2010**, *132*, 8372-8377.
31. Saito, Y.; Motegi, K.; Bag, S. S.; Saito, I., *Bioorg. Med. Chem.*, **2008**, *16 (1)*, 107-113.

Chapter 4: Studies on Overhang and Stem-Loop Sequences

4.1 Introduction

After researching the photodimerisation reaction of anthracene applied to duplex DNA samples and observing the gated behaviour of this system, it was decided to investigate new ways in which to achieve the 'photo-release' of an oligonucleotide. The concept of being able to separate a DNA duplex into its single stranded components through a photochemical reaction remains a desirable target and could potentially have wide ranging applications, for example in light-triggered DNA delivery applications.

As discussed previously in Chapter 3.8 there are a number of factors that can prevent photodimerisation when in the presence of the complementary sequence, which need to be overcome if 'photo-release' is to be made possible. The first major issue involved is the relative orientation of each anthracene tag once the double helical structure of B-DNA is in place. As the number of bases separating each tag is increased, so does the angle of rotation about the vertical axis, increasing the distance between anthracene groups. In addition, the intercalation of the anthracene groups between neighbouring base pairs can also potentially preclude the availability for photodimerisation. A solution to these issues would involve placing the anthracene tags in a location where they are unlikely to intercalate, whilst still being in proximate positions with sufficient flexibility to initiate photodimerisation. The focus of this chapter is an initial investigation of the strategies, in an attempt to overcome these issues.

4.1.1 Nucleic Acid Overhangs

One possible way to generate a photo-release system is to place the anthracene tags on short, single stranded domains on the end terminals of the oligonucleotide, (Figure 1). Extended overhangs on duplex DNA are most commonly known as 'sticky ends' and the single stranded nature of these overhangs allow for interactions with other species through the unpaired nucleotides. This property is important in a number of techniques in biotechnology such as sequence displacement, enzymatic ligation and PCR.[1,2]

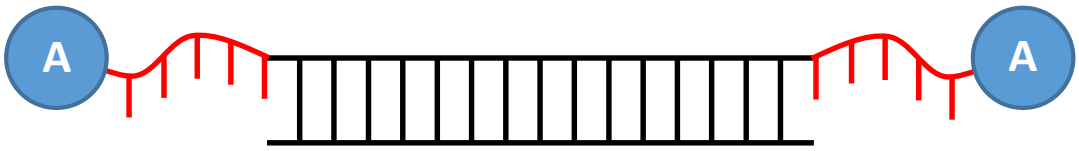


Figure 1. A new sequence design for strategies towards the 'photo-release' of complementary oligonucleotides. The anthracene (A) is placed on short single stranded domains (red) placed on opposite termini of the desired sequence.

Within this position, the anthracene groups would be less likely to intercalate due to the lack of neighbouring base pairs with which it could effectively π-stack, leading to an increased chance of anthracene photodimerisation. It was also anticipated that due to the fact that the overhang segments were unbound it would allow a degree of flexibility and movement that would improve the proximity of the appended anthracene tags despite being bound to the target sequence. Therefore any change in structure due to successful dimerisation of the anthracene tags could lead to an alteration in the binding strength towards the target sequence. The nature of solid phase DNA synthesis would allow the composition of these single stranded domains to be altered and the length increased through the introduction of additional nucleotides. Therefore the flexibility of the

terminal anthracene groups could also be altered to further improve proximity and optimise conditions for photodimerisation.

4.1.2 The Stem-Loop Structure

Whilst the use of a random or a repeating DNA sequence (e.g. poly(A) or poly(T)) within the overhang domains may produce the desired effect of in-*situ* photodimerisation, introducing two complementary overhangs could also lead to the generation of a stem loop. The stem-loop structure results from an intramolecular hybridisation between complementary regions within a given sequence, which allows an oligonucleotide to fold back on itself to form a double helix structure terminating in an unpaired loop, (Figure 2).[3] The formation of a stem-loop is dependent on the overall stability of the intramolecular helix and loop components of the structure. Firstly the helix must be of a sufficient length, typically between 4 and 8 bases, and should have a higher content of G/C base pairs.[4] This provides a greater number of hydrogen bonds and therefore a more stable duplex despite the short length. However the complementary segments of the sequence must remain shorter than that of the loop, which prevents it from competing with the formation of intermolecular duplexes.

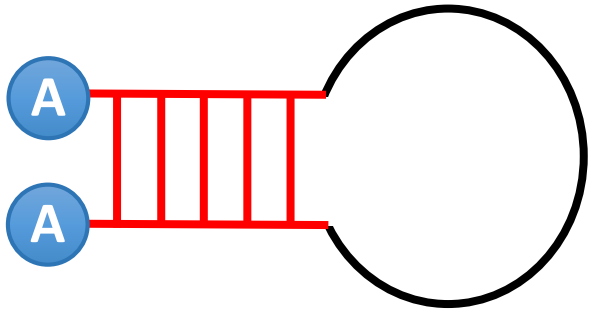
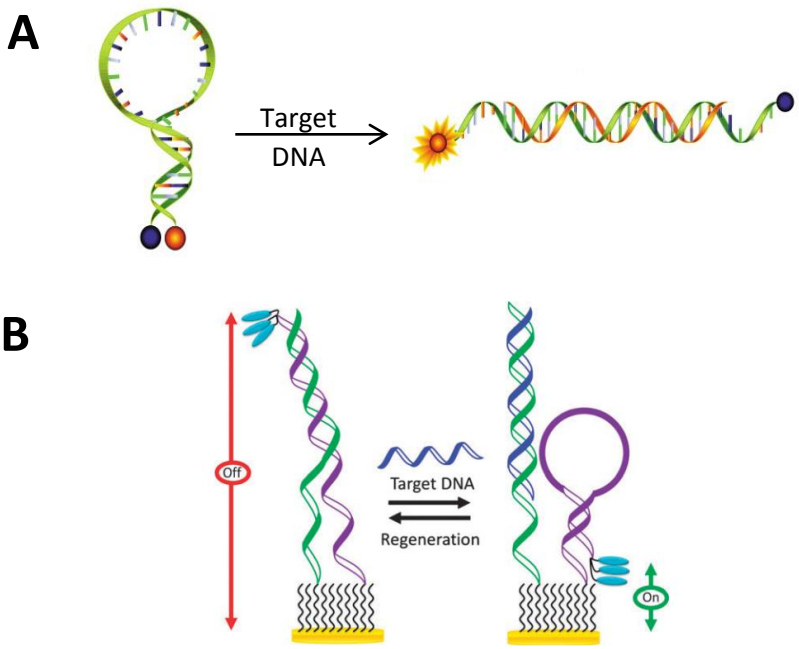


Figure 2. Example of an anthracene modified stem-loop structure. Complementary sequences within the overhang domains (red) lead to intra-strand hybridisation, which could potentially aid the photodimerisation of the appended anthracene groups (A).

Stem-loop sequences have long been utilised in a technique that is used to detect the presence of specific nucleic acids in solution, in a role for which they are more commonly known as a molecular beacon (MB).[5] Molecular beacon probes have a similar design to the stem-loop with the exception of two modifications; Firstly the MB is labelled with a fluorophore on the 5' terminal and a quencher group on the 3' terminal.[6] Secondly the base composition of the loop segment is designed to be complementary to a desired target sequence. The fluorophore and quencher groups within the MB are specifically chosen so that in close proximity, any fluorescence is quenched through resonant energy transfer due to the overlapping emission and absorption spectra of the two chromophores. Therefore in isolation the MB remains in the folded stem-loop structure and no fluorescence is observed. In the presence of a fully matching target strand, hybridisation between the loop and target occurs causing the dissociation of the stem as the new duplex fully extends, (Scheme 1A). This process is initiated due to the fact that the MB sequence can form additional base pairs with the target strand to form a more stable duplex. The resulting shift in duplex structure results in a change in the proximity between the fluorophore and quencher groups, so therefore the presence of the target strand can be detected through an increase in the fluorescent emission.[7]



Scheme 1. (A) An optical MB probe that detects the presence of a target sequence through a change in the fluorescent emission. (B) A similar MB probe that detects the target through an electrochemical response. Image taken from ref. 12.

Molecular beacons have been further developed for use in a variety of applications including a range of assays, PCR quantification and even SNP detection.[8,9,10] Modification of the MB sequence has also allowed the detection of target DNA through an electrochemical response, which can be advantageous due to operational convenience. This technique requires the MB to be bound to an electrode and functionalised with a redox active group such as ferrocene. Binding of the target strand alters the distance between the redox group and the electrode surface resulting in a change in the electrochemical response. This technique was originally developed by Plaxco and further improved by Lai to give the 'signal-on' example as seen in Scheme 1B.[11,12]

The change between stem-loop and intermolecular duplex DNA structures could provide interesting results once the sequence has been modified with anthracene tags. There would certainly be a change in reactivity due to the relative proximity of the anthracene

groups within the two forms. However the success of this strategy is dependent on the initiation of photodimerisation within the intermolecular duplex structure, therefore it would be interesting to observe that any interaction between the anthracene could be promoted by hybridisation of the complementary overhangs, or vice versa.

4.2 Synthesis of a *n* = 7 Anthracene Linker

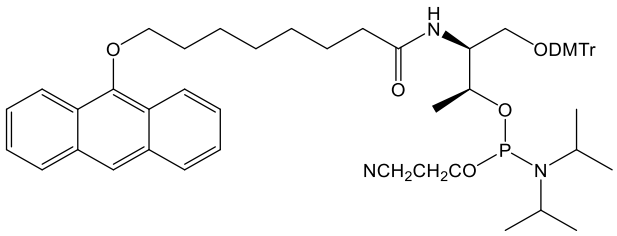


Figure 3. n = 7 anthracene monomer, shown here as the S,S isomer, introducing an extra carbon into the alkyl spacer using the commercially available ethyl 8-bromooctanoate.

Due to the large distances involved in placing the anthracene groups on the extended overhang regions within the sequence it was decided that it was necessary to synthesise an additional anthracene monomer. The new monomer incorporates a longer alkyl spacer with an additional carbon (to give a total of seven carbons) separating the anthracene group and the threoninol diol. The synthesis of the new anthracene monomer was carried out following the same steps as other linker lengths. A new brominated ester (ethyl 8-bromooctanoate) was commercially sourced and successfully coupled with anthrone to provide the initial product in sufficient yield (60%). The synthesis was then continued through the remaining steps to generate the final phosphoramidites (Figure 3), incorporating both D and L isomers of threoninol. It was anticipated that the additional length of this linker would provide greater range and flexibility to the anthracene group, aiding the photodimerisation within this new series of sequences.

4.3 Sequence Design and Strategies

One of the major observations from the previous chapter was that increasing the number of bases separating the anthracene tags also increased the structural alteration induced upon photodimerisation. Therefore an initial sequence was synthesised to investigate the effect of placing the anthracene modifications further apart. Sequence **S17** was synthesised based on the original 15-mer sequence used in Chapter 3 and places the anthracene tags on the 5'and 3' termini of the oligonucleotide (Table 1).

*Table 1. Sequence composition of strands synthesised for photo-release studies. (**X** = anthracene)*

Sequence	Linker	Sequence Composition (5'-3')
S17	7D	**X**TGGACTCTCTCAATG**X**
S18	7D	**X**TTTTGGACTCTCTCAATGTTT**X**
S19	7D	**X**TTTTTGGACTCTCTCAATGTTTT**X**
S20	7D	**X**GCGACTGGACTCTCTCAATGGTCGC**X**
S21	-	GCGACTGGACTCTCTCAATGGTCGC
T0	-	CATTGAGAGAGTCCA
S17-D$_1$	7D	**X**TGGACTCTCTCAATG**X**
S20-D$_1$	7D	**X**GCGACTGGACTCTCTCAATGGTCGC**X**

In solid phase oligonucleotide synthesis each sequence is synthesised in the 3' to 5' direction and typically suppliers provide solid support columns that have been functionalised with the first nucleotide in order to help improve yields (Commercially available with a range of nucleotides). Therefore in order to append an anthracene modification to the 3' terminal a different solid support column is used containing only a DMTr protected hydroxyl group bound to the solid support, also known as a phosphate column, Figure 4. This allows the anthracene monomer to be appended to the oligonucleotide to the 3' end with no preceding nucleotide, however cleavage from the solid support leaves the 3' phosphate group in place. This means that any sequence

synthesised using this technique will contain an extra phosphate group compared to those made through conventional synthesis.

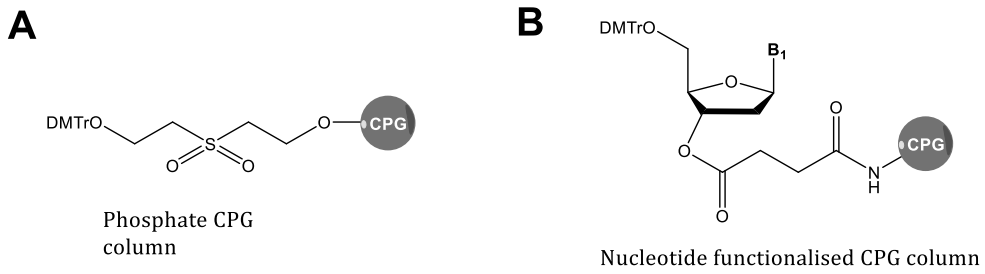


*Figure 4. Comparison of commercially available CPG synthesis columns. The phosphate column (**A**) contains simply a free hydroxyl group, which allows the user to select any chemically modified monomer as the initial 3'-nucleotide. Conventional solid supports (**B**) are functionalised with the first nucleotide of the sequence.*

Sequence **S17** does not require any base substitution in order to introduce the anthracene tags, and at the 5' and 3' terminal positions, the anthracene tags are less likely to be in an intercalated position due to lower π-stacking efficiency at the end of the duplex compared to when the tags are placed with bases either side (Chapter 3). This sequence also utilises the longer *n* = 7 anthracene linker, which could potentially provide the extra flexibility and length to allow the anthracene groups to photodimerise, despite being at opposite ends of the oligonucleotide. Therefore the additional linker length combined with a reduction in π-stacking at the terminal positions could lead to the possibility of photodimerisation of the anthracene tags in the presence of the complementary sequence, and so were included in the aforementioned in-*situ* irradiation experiments.

A second batch of sequences were also developed based on the same principle of placing the anthracene tags on the terminal positions of the oligonucleotide. However in this case sequences **S18**, **S19**and **S20** extend the sequence past the normal-15 mer sequence to

include the short overhang segments (Table 1). Sequences **S18** and **S19** utilise a poly(T) overhang of varying lengths (3 and 4 thymines respectively), whereas **S20** includes two complementary overhang segments (5 base pairs in length) in order to investigate stem-loop formation. The complementary overhangs in sequence **S20** were designed with a high G/C content in order to form a stable stem-loop structure. Sequence **S21**, a non-modified version of the stem-loop sequence was also synthesised as a control.

4.4 Results and Discussion

4.4.1 Stem-Loop Fluorescence Analysis

An important aspect of the stem-loop structure is that upon formation the two terminal ends of the oligonucleotide are brought into close proximity. As discussed previously this property forms the basis of many sensing applications based on molecular beacons, which utilise a stem-loop sequence modified with fluorescent reporter groups to detect the presence of complementary sequences. Anthracene can also give a varying fluorescent response depending on the proximity of a secondary anthracene group, which gives rise to the formation of anthracene excimers.[13] Sequence **S20** was tested through fluorescent analysis to determine whether the formation of the stem-loop structure promoted the formation of excimers between the two terminal anthracene groups. Figure 5 displays the normalised fluorescent spectra of sequence **S20** with and without the presence of the complementary strand **T0**. The presence of excimers is

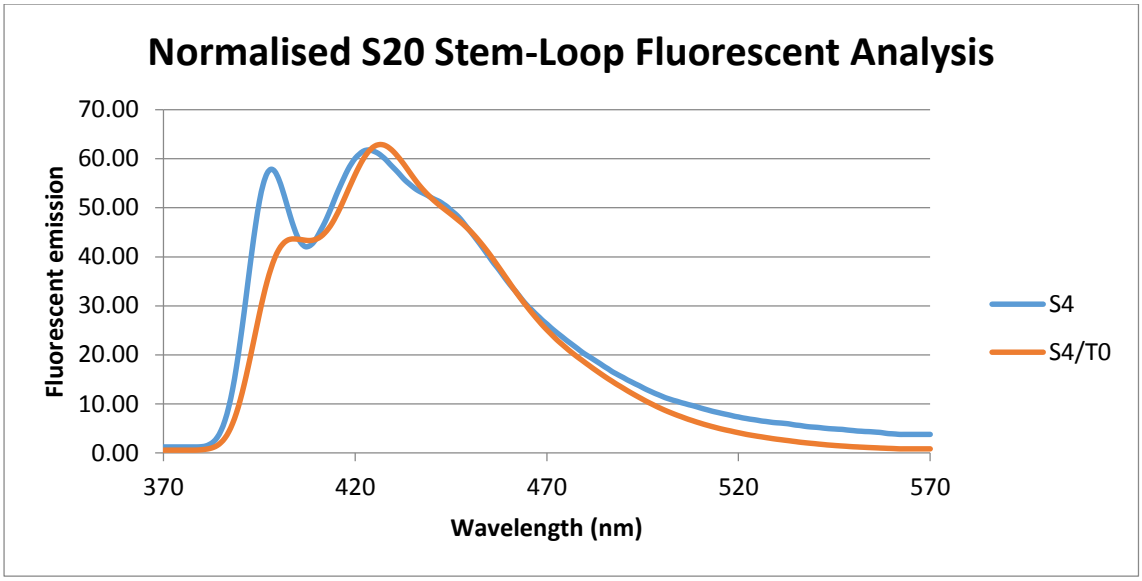


*Figure 5. Normalised fluorescent spectra of **S20** with and without the presence of the complementary strand **T0**. The lack of increase in longer wavelength emission despite their proximity indicates the anthracene tags are π-stacking with flanking bases and are not available for excimer formation.*

detected by an increase in the fluorescence at the longer wavelengths between 470 and 570 nm. However the increase in longer wavelength emission detected when **S4** was analysed was not of the magnitude that would have been expected based on earlier research of anthracene excimer fluorescent emission. This may indicate that either the longer n = 7 linker used provides too much flexibility and reduces the probability of contact between the anthracene groups or that the anthracene groups have a preference to π-stack with the terminal nucleobases leading to a reduction in excimer formation.

4.4.2 Irradiation

The first task in assessing these newly synthesised sequences was to test their effectiveness to photodimerise whilst in the presence of the complementary strand. In order to do this a sample of each sequence was prepared with an equimolar concentration of the target sequence **T0** in a degassed, buffered solution and annealed

prior to testing. Each sample was then irradiated with UV light (365 nm) and monitored through UV spectroscopy to measure the disappearance of the anthracene band (374 nm). Figure 6 displays the resulting reaction profiles, measuring the change in absorbance (ΔA %) against time as the reaction progressed. As can be seen from the data there was a stark difference between the reactivity of the different sequences within this experiment. Sequences **S18** and **S19** displayed no reactivity when exposed to UV light, indicating that the poly(T) composition of the single stranded overhangs may actually have been a hindrance towards photodimerisation. In contrast the stem-loop sequence **S20** readily dimerised under the same conditions, reaching almost a 90% conversion after 1 hour of irradiation. Therefore clearly a change to complementary base-pairing in the overhangs favours intramolecular photodimerisation, presumably through a combination of facilitating the close proximity of the reacting anthracenes and stabilising the resulting photoproduct.

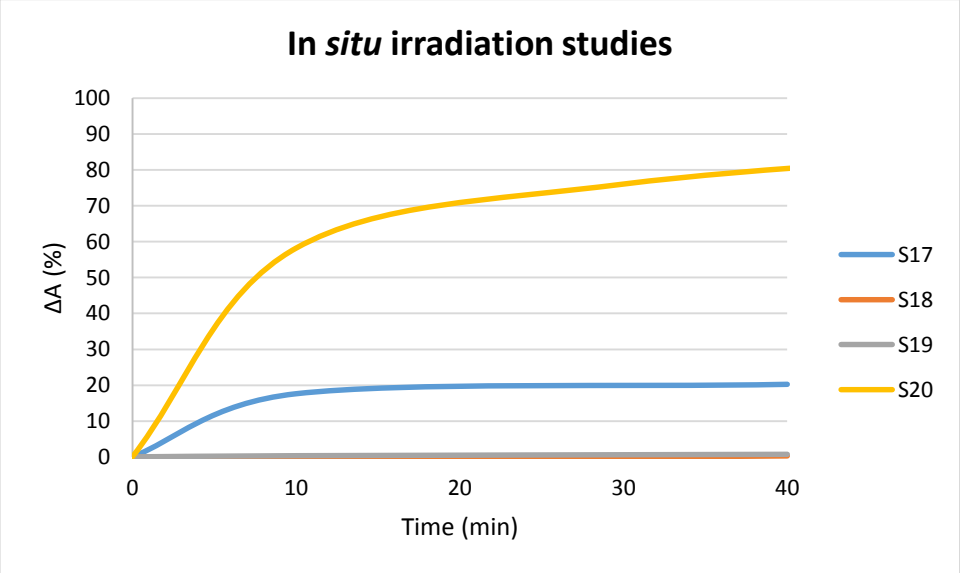


*Figure 6. Samples **S17**, **S18**, **S19** and **S20** are hybridised with the complementary strand **T0** and irradiated as duplex DNA. This graph presents the UV absorbance measurements (365 nm) used to monitor the progress of the reaction over the course of 40 mins, giving an indication of sample reactivity. Only **S20** displays significant reactivity when irradiated.*

It was also interesting to see that sequence **S17** also displayed a degree of reactivity upon irradiation, obtaining a 20% conversion after 40 minutes of irradiation. This result is good evidence for intercalation contributing towards the suppression of photodimerisation when irradiated in-*situ*. With minimal π-stacking interactions between the anthracene groups and the bases within the oligonucleotide there should be a greater availability for photodimerisation. Therefore despite the fact that the anthracene groups have been separated by a much larger distance (15 base pairs) this system has shown greater reactivity than **S15** and **S16** (Chapter 3) under the same conditions, which have much smaller tag separations (3 and 5 bases respectively). However sequence **S17** has not demonstrated a similar reactivity to that of **S20**, indicating that it is the additional interaction afforded by the complementary, single stranded overhangs that is crucial for photodimerisation in-*situ*.

4.4.3 Control Studies to Determine the Effect of Weak Points Introduced by Anthracene Modification.

As part of these studies sequence **S17** was also irradiated in the single stranded state and the D_1 photodimer was isolated through preparative HPLC. When the isolated photodimer was tested through T_m analysis the data obtained provided some interesting findings and offered some insight on earlier analyses. Table 2 lists the T_m data for **S17** in the form of both the starting material (SM) and the photodimer (D_1). The first major observation from this data is that the SM T_m (57 °C) is comparable with that of the unmodified **S0/T0** duplex (55 °C). Owing to the fact that modification of **S17** does not involve substitution of any of the bases within the 15-mer sequence, there is no loss in hydrogen bonding and hence a comparable result. It was however surprising to observe

a T_m transition at 27 °C for **S17-D$_1$**, when sequence **S16**, which has a smaller base separation, displays no hybridisation towards **T0** and exhibits an ON/OFF behaviour upon photodimerisation (Chapter 3). This goes against the trend that increasing the base separation also increases the ΔT_m value upon dimerisation and suggests that there may be more factors involved for this system, which impacts upon the previous results described in Chapter 3.

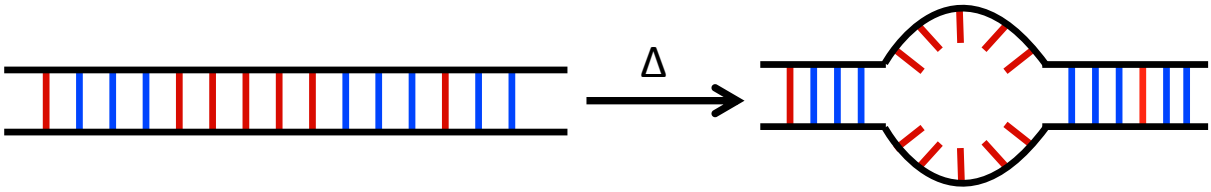
*Table 2. T_m data for sequence **S17**. (Data obtained from sample irradiated as a single strand and the photodimer isolated through preparative HPLC, not as a result of photo release) The data suggests duplex binding despite a very large base separation, which was unexpected when compared to T_m data from **S16**.*

	S17
T_m (SM)	57 °C
T_m (D_1)	27 °C
ΔT_m	-30 °C

S17: **X**TGGACTCTCTCAATG**X**

A known property of DNA is that sequence composition can have an effect on the resulting T_m stability of the duplex structure. For example, two sequences with the same number of G/C and A/T base pairs can have different melting temperatures, depending on the relative arrangement of those base pairs within the oligonucleotide. An accumulation of A/T pairs within a particular segment of the sequence can lead to weak points due to the lower number of hydrogen bonds within the A/T pair. This leads to what has been dubbed as 'pockets' within the duplex as the A/T pairs dissociate at lower temperatures, (Scheme 2).[14,15] These pockets created by the weak points can destabilise the remainder of the duplex structure, leading to a lower overall thermal stability compared to a more random arrangement of the nucleobases. This effect provides subtle differences in the T_m values as demonstrated by Table 3, which displays data generated from T_m prediction software

from two sequences with the same ratio of A/T to G/C pairs (7:8).[16] Whilst **A** has a more random arrangement, **B** includes a weak point at the centre of the sequence and therefore the T_m decreases by 3.4 °C.



Scheme 2. An accumulation of A/T pairs (red) within a sequence can lead to weak points that create pockets upon heating. This pocket can further destabilise the remaining structure, lowering the overall T_m.

*Table 3. T_m data from prediction software that demonstrates the change in T_m when weak points are present. Sample **A** & **B** contain the same no. of A/T and G/C base pairs, but whilst sample **A** has a random arrangement of base pairs sample **B** introduces a weak point of A/T pairs highlighted in red. Calculated using online software (ref. 16), 5μM oligo concentration, 100 mM NaCl.*

Sample	Sequence	T_m (°C)
A	5′-TGGCAGTTCGTGACA-3′ 5′-ACCGTCAAGCACTGT-3′	60.8
B	5′-GGCG**TATTTAA**CGGC-3′ 5′-CCGC**ATAAATT**GCCG-3′	57.4

This behaviour could play an important part within this research on modified oligonucleotides, as the anthracene tags, with no ability to hydrogen bond, can also introduce points of weakness into the duplex structure, leading to the formation of pockets. Prior to irradiation, anthracene intercalation can compensate for the loss in hydrogen bonding within the starting material through π-stacking. However in the form of the photodimer these points of weakness can contribute towards the ΔT_m in addition to any structural change induced by photodimerisation. The extent to which these weak-points destabilise the overall duplex structure is dependent on the length and composition of the sequence on either side of the weak point. Longer sequences with higher G/C content will be able to better withstand the removal of hydrogen bonds.

Therefore it is important to investigate the relationship between the position of the anthracene tags and the how this can affect duplex stability.

Table 4 contains T_m data for the double C3 modified sequences **S14-C**, **S15-C** and **S16-C** reviewed in Chapter 3.9. This data is subtracted from the T_m of the unmodified duplex **S0/T0** (54 °C) to generate the value α-T_m, which on comparison should give a better understanding of how base separation has an effect on duplex stability through pocketing, without the presence of anthracene π-stacking. In order to compare the data from all three sequences, the α-T_m for sequence **S15-C** must be corrected for the difference in base pair substitution. For this purpose a control sequence for **S15-C** where two A/T pairs have been swapped for G/C (Table 4), was analysed through T_m prediction software to generate a T_m of 61.2 °C (Calculated using ADT Biophysics,[16] using a nearest neighbour algorithm, 5 μM oligo concentration, 100 mM NaCl) . Using this value to generate the α-T_m for **S15-C** it can be evaluated against the values for **S14-C** and **S16-C**.

*Table 4. Comparison of C3 modified sequences with **S0/T0** to observe trends in base separation and the effect of pocketing. *Value adjusted to compensate for difference in base pair substitution.*

Sequence	Sequence Composition	T_m (°C)	α-T_m (°C)
S14-C	TGGACT**Y**T**Y**TCAATG	30	-24
S15-C	TGGAC**Y**CTC**Y**CAATG	31.5	-29.7*
S16-C	TGGA**Y**TCTCT**Y**AATG	19.5	-34.5
S15 Control	TGGAC**C**CTC**C**CAATG	61.2	-

On analysis of the α-T_m data, a trend emerged of an increasing negative value as the anthracene tags are moved further towards the 3' and 5' terminals. It is clear that moving from sequence **S14-C** to **S16-C** the base pair sequence on either side of the C3

modification gets shorter and is less able to withstand the point of weakness introduced by the modification, therefore the α-T_m value increases consecutively. Sequence **S17**, which does not require substitution of any bases on modification, does not experience the effect of 'pocketing' and the change in T_m (Δ) is simply due to a structural alteration induced by photodimerisation. However despite this change in structure the bases can still effectively hydrogen bond to an extent that a T_m transition is observed instead of the ON/OFF behaviour seen with **S16**, therefore showing the importance of the effect of 'pocketing' within these studies and its influence on the T_m results.

4.4.4 Reversion Kinetics

The reversion kinetics of sequence **S17** were also analysed in order to investigate how the larger separation between the anthracene tags effects the stability of the resulting photodimer (Figure 7). It was expected that the larger base separation would increase the amount of structural strain placed on the photodimer, resulting in a faster rate of reversion. However the rate obtained was 2.1×10^{-3} s^{-1}, similar to the results observed for sequences **S14**-**S16**. This result suggests that no additional strain was placed on the photodimer and it is more likely that the flexible single stranded oligonucleotide is capable of folding in such a way that accommodates the anthracene photodimer, keeping it relatively stable.

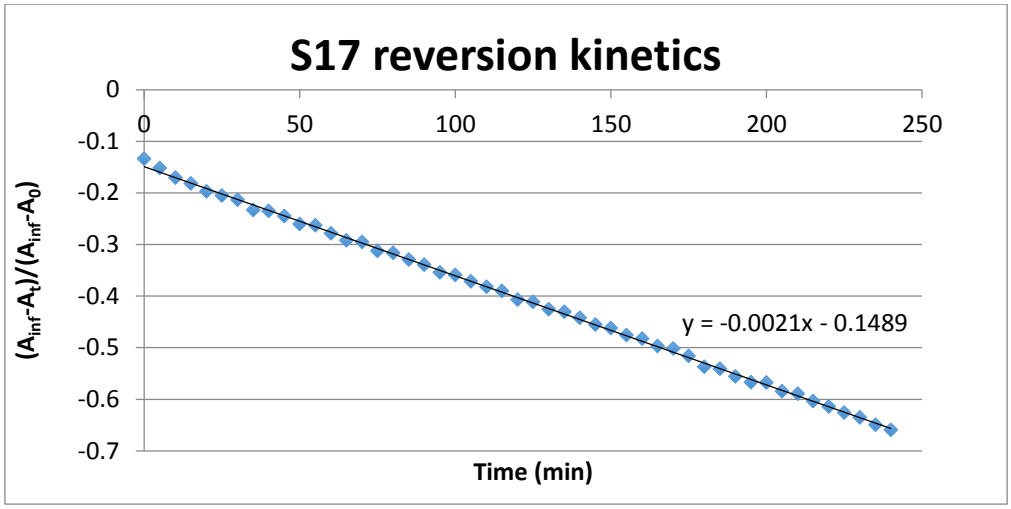


*Figure 7. Reversion data for sequence **S17**, the gradient of this data can be used to calculate the value k_{diss}. (Sample heated to 80 °C and changes in absorbance recorded at 1 minute intervals)*

4.4.5 Stem-Loop T_m Analysis

As the stem-loop sequence **S20** was the only sequence capable of substantial in-*situ* photodimerisation, it was therefore necessary to complete a comprehensive T_m analysis to observe what effect this process had on the hybridisation between **S20** and **T0**. The first step was to analyse **S20** prior to irradiation and compare this data with that of the unmodified sequence **S21**. Figure 8 displays the raw data observed from the VT-UV experiment, whilst Table 5 lists the T_m values obtained for each sample from that data. After analysing Figure 8 it was clear that anthracene modification plays an important part in influencing the stability of the stem-loop structure. Firstly the anthracene modified stem-loop **S14** displayed a 14 °C increase in stability over the unmodified **S15**, indicating that the anthracene groups were able to stabilise the short stem-loop duplex. This could potentially be explained by π-stacking interactions, which prevent fraying of the end nucleobases, providing a stronger duplex.

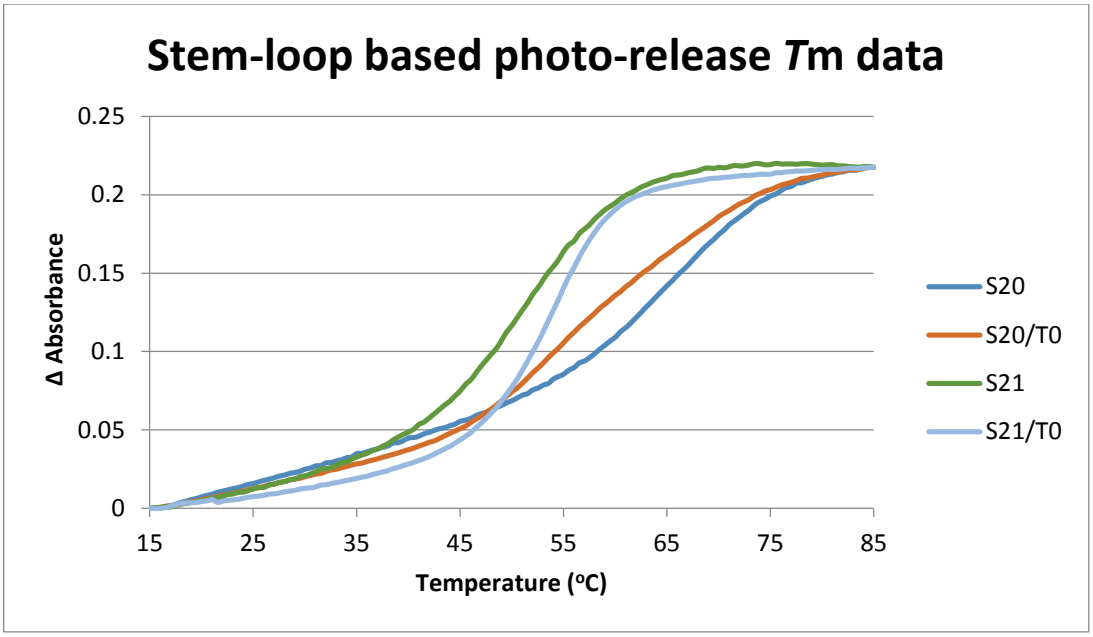


*Figure 8. Raw T_m data comparing the stem-loop sequence **S20** with the unmodified **S21**. The sequences displayed opposite behaviours upon addition of **T0**, with the T_m decreasing for S20 and decreasing for **S21** indicating that anthracene has stabilised the stem loop structure.*

Table 5. T_m values obtained from the VT-UV data.

Sample	T_m (°C)
S20	65
S20/T0	54
S21	51
S21/T0	54
S0/T0	55

Upon analysis of the T_m data it was observed that **S20** and **S21** displayed opposite behaviours upon the addition of the target strand **T0**. Whereas the T_m for **S21** increased by 3 °C upon duplex formation, consistent with the formation of additional base pairs, the T_m for **S20** decreased by 11 °C. The data suggests that the addition of the target strand does result in the opening of the stem-loop structure and the formation of the duplex, however the subsequent drop in the T_m value implies that this duplex is not as stable due to the loss of the interaction with anthracene and hybridisation with the target strand may be impeded. This result was confirmed on closer inspection of the derivative data

from this experiment (Figure 9), an unsymmetrical curve rather than the normal Gaussian shape was observed.[17] This data indicates that the T_m curve is actually a composite of two transitions from both the stem-loop structure and the **S20**/**T0** duplex, therefore addition of **T0** has not been successful in fully hybridising with **S20** and removing the stem-loop structure. Nonetheless the maximum absorbance gradient appears at the lower temperature of 54 °C, which would suggest there is still significant binding with **T0**. Therefore **S20** may still be a viable target for photo-release if it was used in an excess with the target strand.

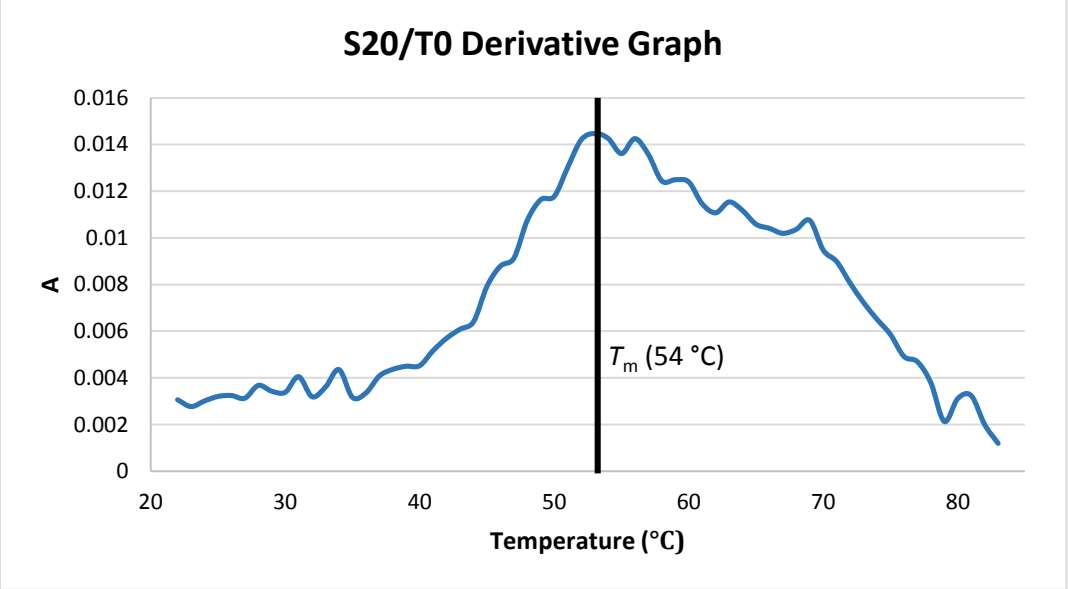


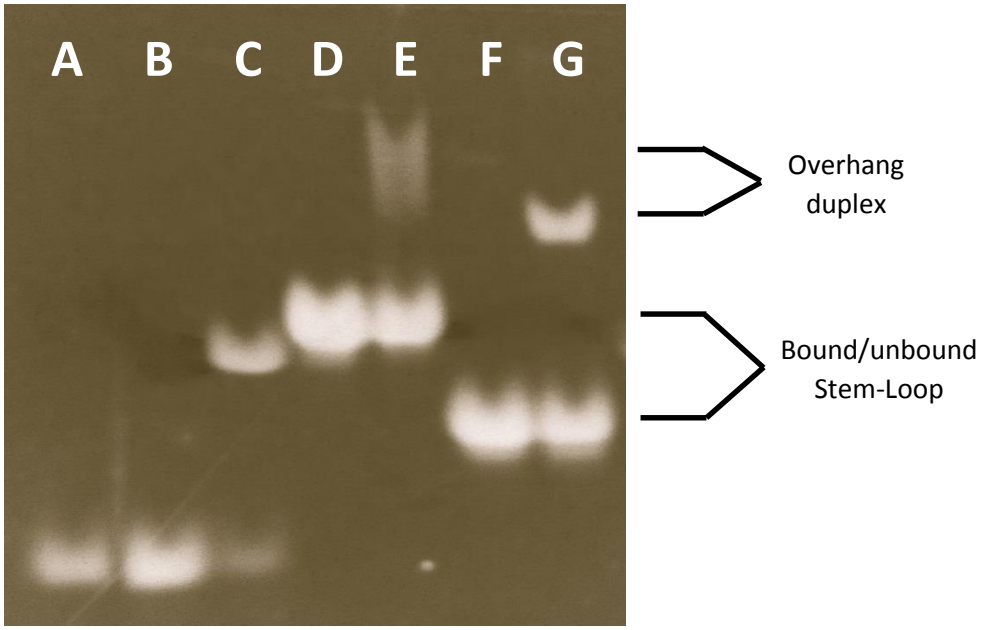
*Figure 9. Derivative of the **S20/T0** T_m transition curve indicating incomplete hybridisation with the target strand.*

A similar observation was made by Inouye and co-workers, who studied the excimer emission properties of a stem-loop oligonucleotide labelled with pyrene on both the 3' and 5' ends. They found that the pyrene modification was able to act as an additional base pair, increasing the stability of the modified oligonucleotide 13 °C over the native DNA sequence when analysed through VT-UV spectroscopy. However, unlike the anthracene modified sequence **S20**, this increase in stability did not appear to disrupt the binding

towards the target DNA and the stem-loop was successfully opened up upon addition of the complementary sequence as observed through changes in the fluorescence spectra.[18]

To provide further evidence of the binding between **S20** and **T0,** a native GE experiment was implemented that utilised different combinations of the two sequences to compare differences in migration speed. Figure 10 displays the resulting UV shadow image of the GE experiment and the various sample lanes are detailed in the associated table. Lane **D** contains the single stranded **S20** sample which under the given conditions forms the stem-loop secondary structure. As can be seen from this image the formation of the stem-loop structure has slowed migration through the gel to a rate comparable to that of the **S0**/**T0** control duplex. Upon addition of one equivalent of the target within lane **E,** a new band appears confirming the interaction between **S20** and **T0** where the stem-loop structure has been removed. The migration speed of this duplex seems to have been further hindered by the overhangs within the sequence which can interact with the gel, a result that is consistent with observations from the unmodified sequence **S21** in lanes **F** and **G**. The smudged and less intense appearance of the **S20**/**T0** band in lane **E** evidence of the weaker binding of this duplex compared to the stem-loop, whereas the band for the **S21**/**T0** duplex appears well defined. A major observation from this image however was that there is no appearance of a band representing unbound **T0** within lanes **E** or **G** despite the stem-loop structures being present. One explanation for this result is that sequences **S20** and **S21** bind **T0** forming conventional duplexes, although due to the long length of the loop segments they can also do so whilst forming their stem-loop structures. This result was not expected, but nevertheless the analysis was taken further to

investigate what effect anthracene photodimerisation had within this system, as detailed below.



Lane	A	B	C	D	E	F	G
Sample	**S0**	**T0**	**S0/T0**	**S20**	**S20/T0**	**S21**	**S21/T0**

*Figure 10. Native gel electrophoresis experiment comparing the ability of **S0**, **S20** and **S21** to hybridise with the target strand **T0** (ambient temperature). (20 % native PAGE gel, 50 mM NaCl, 1 x TB buffer, 19 hour run time, 100V)*

4.3.6 Photodimer T_m analysis

A T_m analysis of the **S20** photodimer, **S20-D$_1$**, in the absence and in the presence of the target **T0** was carried out (Figure 11). When the **S20-D$_1$** photodimer was analysed by VT-UV spectroscopy, it was observed that the T_m transition had shifted to a higher temperature range and only the beginning of the transition could be seen between 70 – 85 °C.

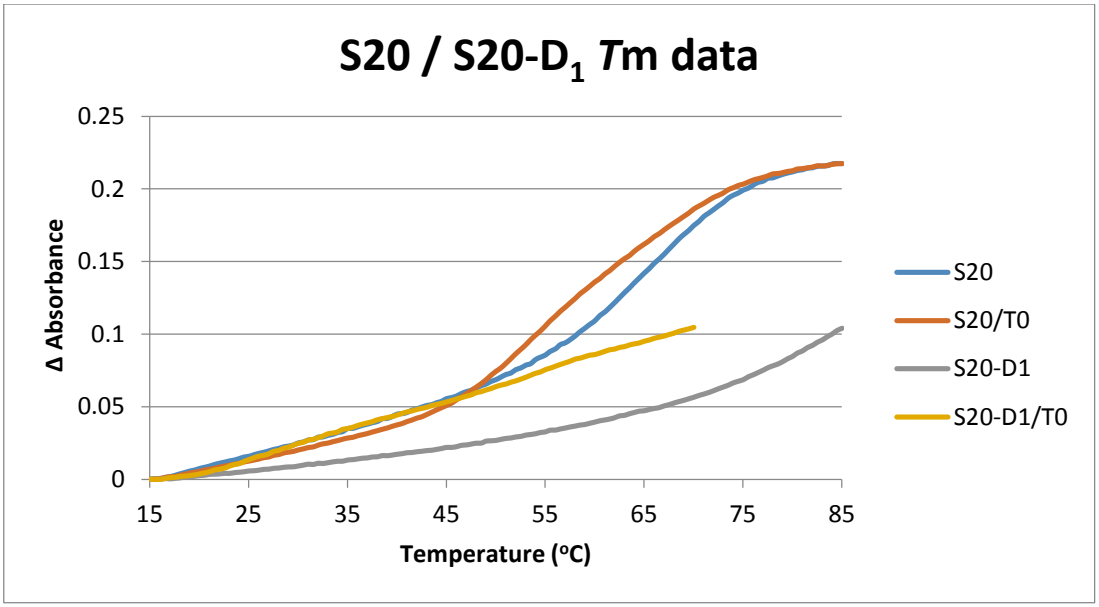


*Figure 11. Melting curve (T_m) analysis of the stem-loop photodimer **S20-D₁**. (Note: temperature range of **S20-D₁/T0** is restricted to 70 °C to limit the effects of photodimer reversion.) The T_m for the **S20-D₁** photodimer appears to be off the scale, indicating that the dimer has stabilised the stem duplex.*

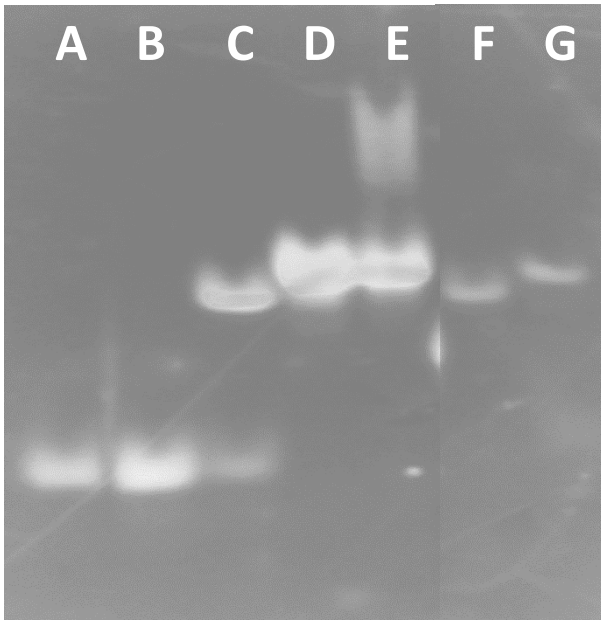
Table 6. T_m values obtained from the VT-UV data.

Sample	T_m (°C)
S20	65
S20/T0	54
S20-D₁	>75
S20-D₁/T0	-

When the anthracene tags are photodimerised within the stem-loop structure of **S20** they covalently link the 3' and 5' ends of the oligonucleotide. The covalent linkage has the effect of stabilising the hydrogen bonding within the stem duplex, holding the ends of the oligonucleotide together in such a way that the nucleotides can only dissociate upon greater thermal heating compared to the non-dimerised sample **S20**.[19] However, upon extensive heating of the sample the covalent linkage within the oligonucleotide is removed and therefore the normal T_m of the anthracene stem-loop is restored. When the **S20-D**$_1$ photodimer was paired with the target strand **T0** a similar observation was made,

with no observable T_m. However an increase in the change in absorbance compared with **S20-D$_1$** indicates that there may still be some residual binding between the unpaired bases within the loop structure and the target. Deviation from the normal sigmoidal curve observed for the non-dimerised **S20/T0** is good evidence that the covalent linking of the end terminals within the stem-loop through anthracene photodimerisation has altered the binding behaviour towards the target **T0**. This is a promising result in terms of achieving photo-release, however further analysis was needed to confirm to what extent the **S20-D$_1$** could potentially bind to the target, as outlined below.

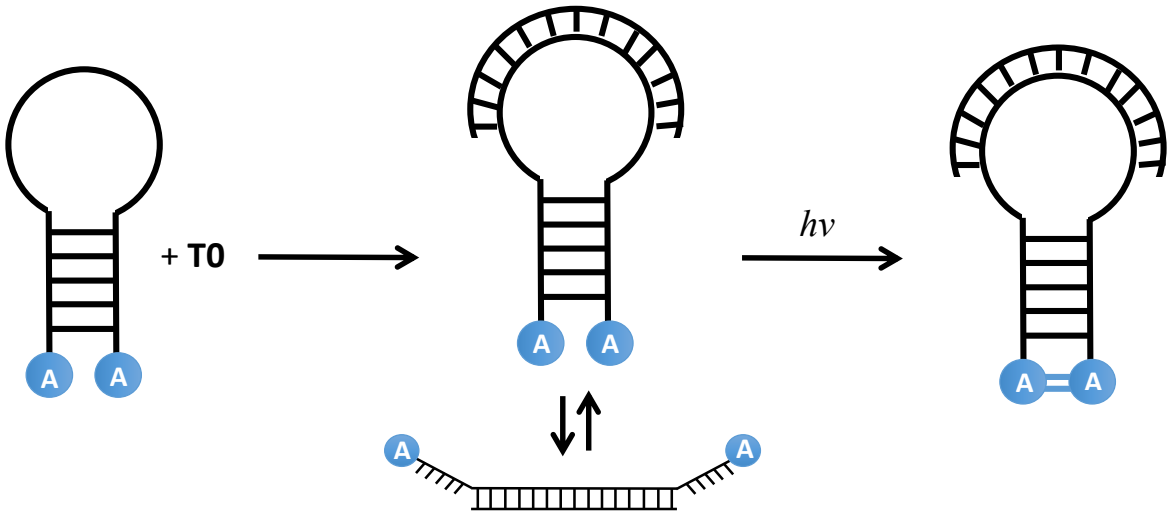
Following results from T_m analysis a native gel electrophoresis experiment was carried out to determine whether the technique could confirm the presence of binding between the **S20-D$_1$** photodimer and **T0**. Figure 12 displays the resulting UV shadow image from this experiment. Lanes **F** and **G** contain the samples **S20-D$_1$** and **S20-D$_1$/T0** respectively and it was observed from the gel image that these samples behaved in a similar fashion, migrating through the gel at relatively the same speed. This result is a good indication that the strengthening of the stem duplex through photodimerisation has prevented **T0** from opening the stem-loop structure to form a conventional duplex. However there was no indication of the release of **T0**, with no appearance of a secondary band that could be compared to the **T0** sample within lane **B**. Therefore it is most likely that the target strand **T0** is still bound to **S20-D$_1$** within the stem-loop structure and therefore the desired photo-release has not occurred.



Lane	A	B	C	D	E	F	G
Sample	**S0**	**T0**	**S0/T0**	**S20**	**S20/T0**	**S20-D$_1$**	**S20-D$_1$/T0**

*Figure 12. Native gel electrophoresis experiment used to determine the effect of anthracene dimerisation towards **S20/T0** binding. (20 % native PAGE gel, 50 mM NaCl, 1 x TB buffer, 19 hour run time, 100V)*

In summary, gel electrophoresis and T_m results have indicated that modification of the stem-loop sequence with anthracene on the terminal positions have increased the stability of the stem duplex through π-stacking, acting as an additional base pair. This increase in stability makes it more difficult for the complementary sequence **T0** to fully hybridise and open the stem-loop into an intermolecular duplex. Although binding and opening of the stem-loop does occur, this happens only partially and the system lies in an equilibrium between the open and closed states. Photodimerisation of the terminal anthracene groups locks the stem-loop in the closed structure, however the complementary sequence can still pair with the bases within the loop of sequence, therefore 'photo-release' has not been achieved (Scheme 3).



Scheme 3. Summary of the observations from gel electrophoresis and T_m data of the anthracene (A) modified stem-loop sequence. Addition of the complementary sequence results in an equilibrium between the open and closed states of the stem-loop structure.

4.5 Competing Strands and FRET Modification

Although it appeared from the previous gel electrophoresis and T_m experiments that **S20** had provided a negative result in terms of 'photo-release', through further experimentation it could still be possible to achieve this elusive property, albeit in a different form. It is clear that the target **T0** would be most stable in the form of the fully hybridised duplex with **S20**, and from the gel electrophoresis in Figure 10 it was observed that it was partially capable of achieving this. The main challenge was overcoming the hydrogen bonding and anthracene interactions that are present within the stem-loop in order to unfold the structure and obtain full hybridisation.

Based on results from the gel electrophoresis in Figure 12 it is known that photodimerisation of the terminal anthracene groups completely inhibits **T0** from unfolding the stem-loop, although binding between the two oligonucleotides is still present. This change in behaviour may also represent a change in how strongly **T0** is

bound to the **S20-D$_1$** photodimer now that it is solely in the form of the stem-loop, rather than in an equilibrium with the open form. The main issue was attempting to utilise any change in binding strength to then progress towards 'photo-release'. With inspiration from the concept of sequence displacement, a series of competing oligonucleotides were developed, which were specifically designed with base pair mismatches so that they would have poorer binding with the target **T0**. The idea being that the mismatched oligonucleotides would only be capable of hybridising with **T0** upon lowering of the **S20/T0** binding strength through photodimerisation.[20]

A number of mismatched sequences were investigated through T_m prediction software in order to decide which sequence would give a suitable T_m that did not exceed the value obtained from experimentation of the **S20/T0** complex (54 °C). Two competing strands were therefore developed, one containing a single C/A mismatch (**S22**) and a second containing two mismatches (**S23**), (Table 7). The prediction software (IDT Biophysics[16]) determined that the **S22/T0** and **S23/T0** duplexes would give a T_m of 44.19 and 30.31 °C respectively and it was anticipated that one of these sequences would be capable of competing for the target **T0**, but only upon anthracene photodimerisation. The use of prediction software to determine the effect of mismatches has been shown to provide a decent level of accuracy, with predicted values within 1.6 °C of the observed values on average, however as the number of mismatches increase so too does the inaccuracy.[16] Therefore these values were only used as a rough guide to decide on the best mismatch sequences for use in this study.

*Table 7. Sequence composition of target and competing oligonucleotides modified with Cy3 (**Y**) and Cy5 (**Z**) FRET chromophores, along with their predicted T_m values (5 μM oligo concentration, 100 mM NaCl, 10 mM phosphate buffer pH 7.0) The location of the C/A mismatches are indicated below.*

Sequence	Sequence Composition (5'-3')	Predicted T_m (°C)
T1	CATTGAGAGAGTCCA**Y**	-
S22	TGGACTC**C**CTCAATG**Z**	44.1
S23	TGGAC**C**CTC**C**CAATG**Z**	30.3

T1/S22

TGGACTC C CTCAATGZ
YACCTGAG A GAGTTAC

T1/S22

TGGAC C CTC C CAATGZ
YACCTG A GAG A GTTAC

In order to complete this experiment a method of identifying the change in oligonucleotide pairs would have to be developed. Using existing techniques such as gel electrophoresis and T_m analysis could prove difficult to confirm the change from one duplex to another. Therefore the target and competing oligonucleotides were modified with the fluorescent groups Cy3 and Cy5 respectively (Table 7) in order to utilise a technique known as Förster resonant energy transfer (FRET).[21] FRET occurs when two chromophores with overlapping fluorescent emission and excitation spectrums are able to transfer energy through non-radiative dipole-dipole coupling (Figure 13). This transfer of energy leads to a decrease in the fluorescence intensity of the donor chromophore and an increase in fluorescence intensity from the acceptor. The process is extremely sensitive to small changes in distance and the transfer of energy can only occur when the two chromophores are within a certain range (typically 10-100 Å). This property has therefore already been used within a wide range of applications, as any process that affects the energy transfer rate between chromophores will allow that process to be quantified using the change in fluorescence emission. FRET is known as the

spectroscopic ruler since it can be used to measure very small distances; however in this case all that is required is the detection of binding between **T1** and either of the competing strands.[22] Hybridisation of the two oligonucleotides would bring the two chromophores within the required range for FRET to occur, confirming the event. The Cy3 and Cy5 dyes can be coupled with the oligonucleotide during solid phase synthesis and is incorporated as a cyanine phosphoramidite. However the dyes are sensitive to the reagents used in some synthesis steps, therefore milder versions of each reagent is used and the coupling procedure is modified in order to protect the dyes. In addition to this damage to the dyes is minimised by preferentially placing them on the 3' terminus of the oligonucleotide, therefore it is the last coupling that is performed and only has to undergo a single synthesis cycle.

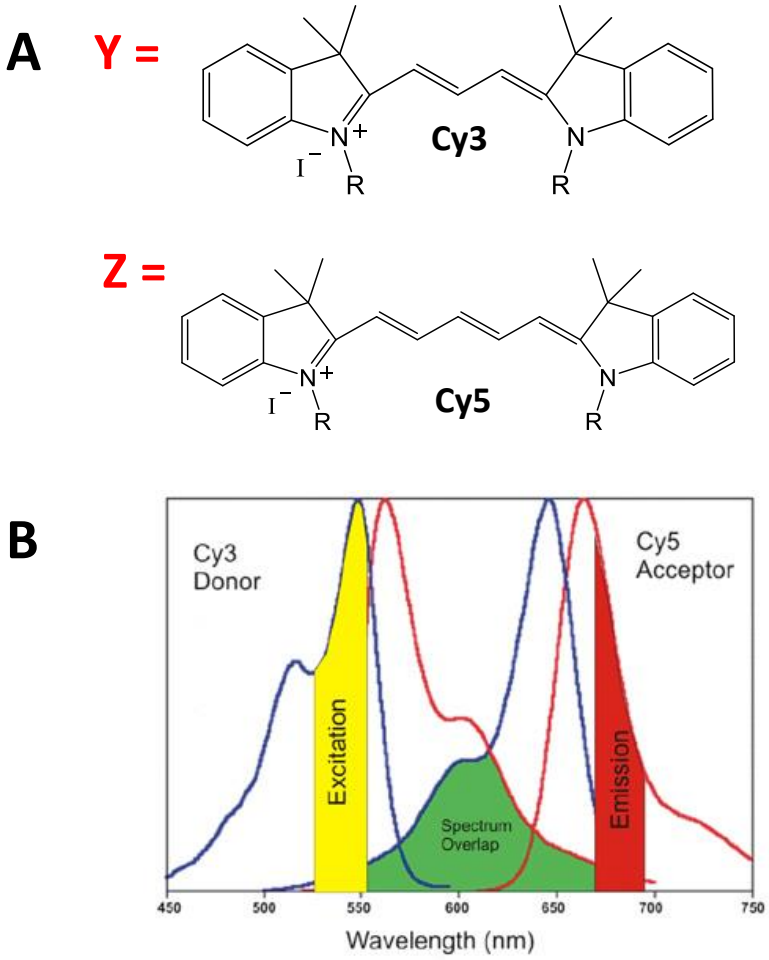
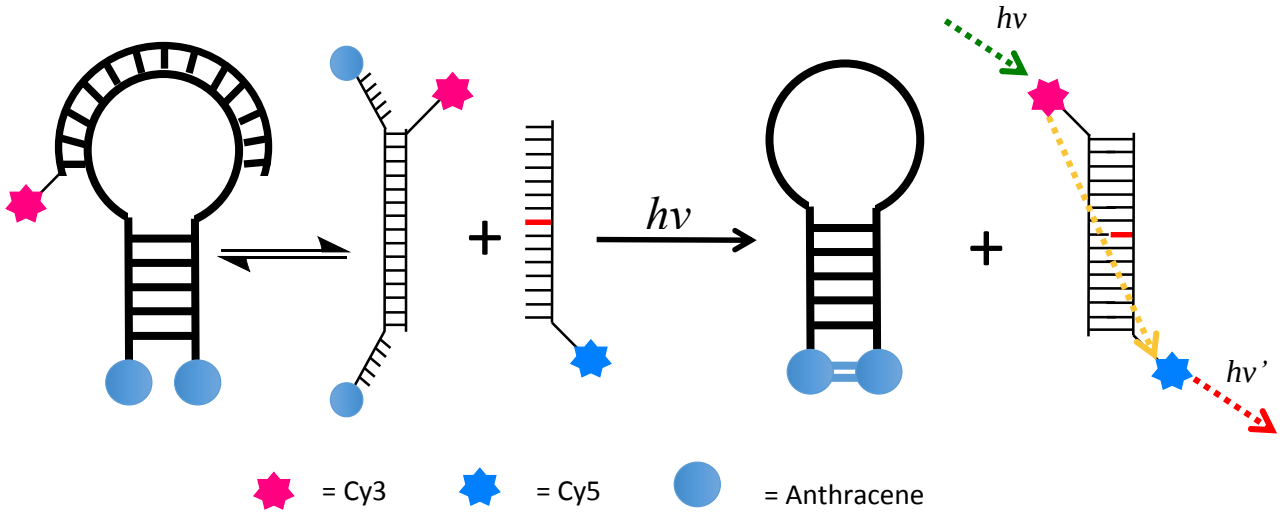


*Figure 13. (A) Structures of the Cy3 and Cy5 FRET pair used in sequences **T1**, **S22** and **S23**. (B) Overlap in the fluorescent excitation and emission spectra allow energy transfer to occur.*[23]

Scheme 4 displays a schematic representation of the desired transfer of **T1** from the stem-loop complex to form a duplex with a competing strand upon photodimerisation. As previously indicated from T_m and gel electrophoresis results, it is likely that the interaction between **S20** and **T0** leads to a mixture of two forms; one where the sequences are fully hybridised in a conventional duplex structure and the other where **T0** is bound to the stem-loop. Therefore the competing strand may be capable of binding with the target from the stem-loop before irradiation, depending on the binding strength between the two strands through the inherent mismatches. However the main challenge is to observe a subsequent increase in binding upon photodimerisation as the remaining

T1 is released. Following photodimerisation it is anticipated that any additional binding between the competing strand and **T1** would result in an increase in FRET observed through a change in the Cy5 emission, (Scheme X).



*Scheme 4. A schematic representation of the displacement of **T1** by the competing strands **S22** or **S23** upon anthracene photodimerisation. The resulting duplex can then be detected through FRET emission.*

Before any experimentation could take place, the parameters of the fluorometer had to be calibrated in order to achieve optimal results. The λ_{max} of the Cy3 chromophore is 550 nm, however due to the overlap of fluorescent spectra Cy5 is also partially excited at this wavelength, (Figure 14). Therefore exciting at this wavelength would produce an emission spectrum displaying the emission from both chromophores, potentially giving a false positive. To overcome this problem the competing strand bearing the Cy5 acceptor chromophore was scanned individually before any experiment so that the data could be subtracted from that obtained during the titration.

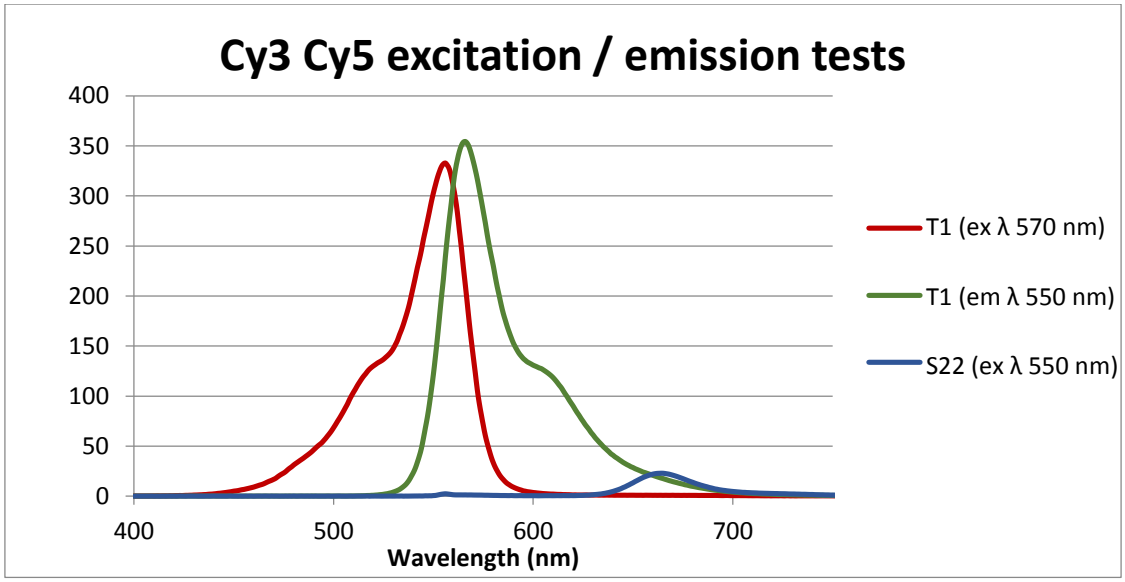
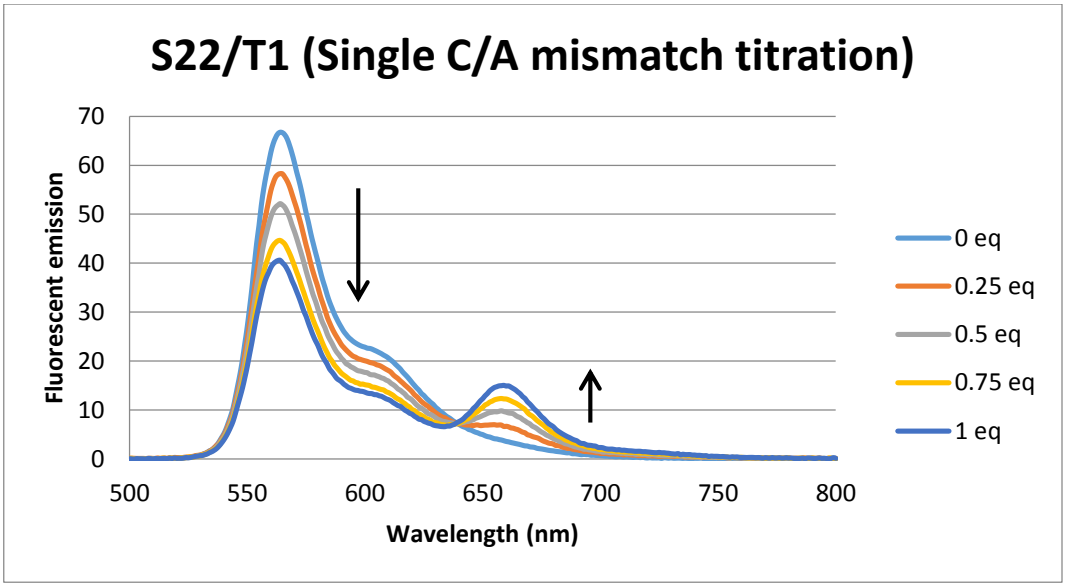


*Figure 14. Excitation (ex) and emission spectra (em) of **T1** and **S22**. Excitation of **T1** at 550 nm can lead to the simultaneous excitation of **S22**, which therefore must be subtracted from the final data.*

With the initial testing of the fluorescently labelled competing strands complete, a trial titration was carried out in order to monitor the change in fluorescence induced by FRET. Figure 15 displays the resulting emission spectra of **T1** when titrated with the single C/A mismatch competing strand **S22**. As the concentration on **S22** is increased there is an observed reduction in the emission at 570 nm with the simultaneous increase of emission at 670 nm through an isosbestic point. The change in emission intensity confirms that hybridisation of the two strands brings the Cy3 and Cy5 chromophores within the required range for FRET to occur.



T1: YACCTGAGAGAGTTAC

S22: XGCGACTGGACTCTCTCAATGGTCGCX

*Figure 15. Fluorescent titration of **T1** against **S22**, displaying a change in the fluorescent emission due to an increase of FRET (after Cy5 subtraction).*

As part of this analysis the T_m of the duplexes between the competing strands and the target **T1** were carried out and correlated with the data given by T_m prediction software. In addition to this, the analysis was run in the presence of the stem-loop sequence **S20** to confirm that the competing strands were not capable of hybridisation with the target **T1** when the stem-loop was in the non-dimerised form. Figure 16 and Table 8 displays the resulting melting curves and T_m data from this analysis, and confirms the lower T_m values for the **T1**/**S22** and **T1**/**S23** duplexes due to the inherent mismatches within these sequences, although not as low as the values predicted through computer software. However in the presence of the stem-loop sequence **S20** both T_m values increased in line with the observations for the interaction between **S20** and **S21**, thus indicating that the competing strands were not capable of fully displacing **T1** from the **S20**/**T1** duplex.

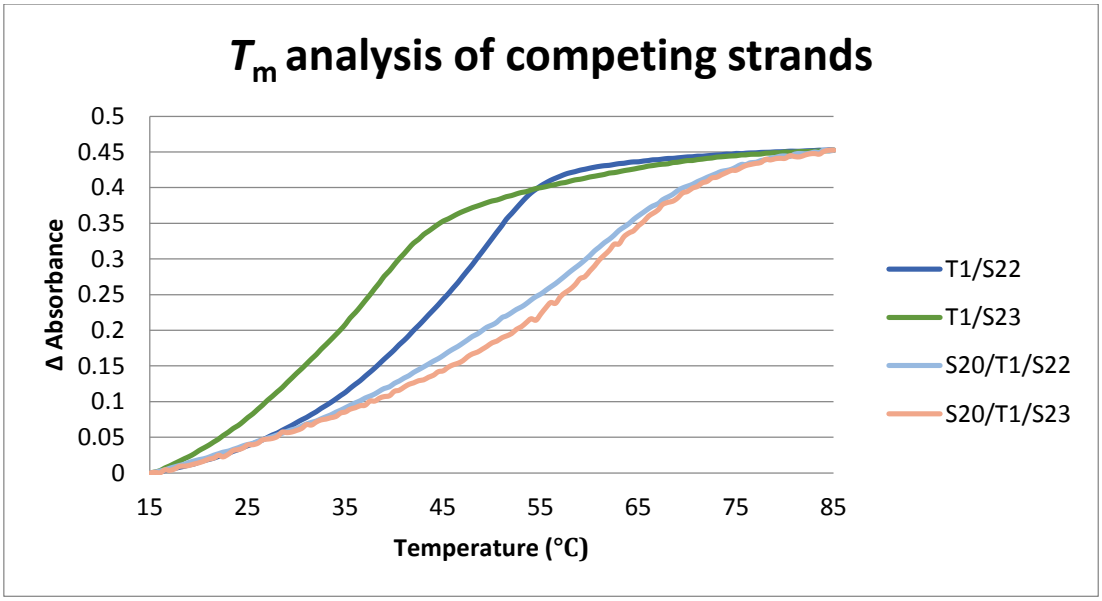


*Figure 16. Melting curve (T_m) analysis of duplexes between **T1** and the competing strands. This data confirms that **S20** binds preferentially to **T1** rather than the competing strands prior to photodimerisation.*

Table 8. T_m values obtained from the VT-UV data.

Sample	T_m (°C)
T1/S22	49.5
T1/S23	39
S20/T1/S22	56
S20/T1/S23	57
S20/T1	54

To test the competing strand strategy, samples were prepared with an equal ratio of **S20**, **T1** and one of the two competing sequences. The samples were made up in a buffered solution, annealed through heating and degassed to remove oxygen before irradiation. It was initially found that the Cy5 dye was sensitive to irradiation and the solutions had to be degassed for an extended period (20 min) before irradiation to avoid degredation. Throughout irradiation the reaction was monitored through UV spectroscopy by measuring the decrease in absorption within the anthracene band, (Figure 17). Results showed successful photodimerisation of the anthracene tags over a period of 180 mins and the irradiation was halted at set intervals in order to obtain a fluorescent scan.

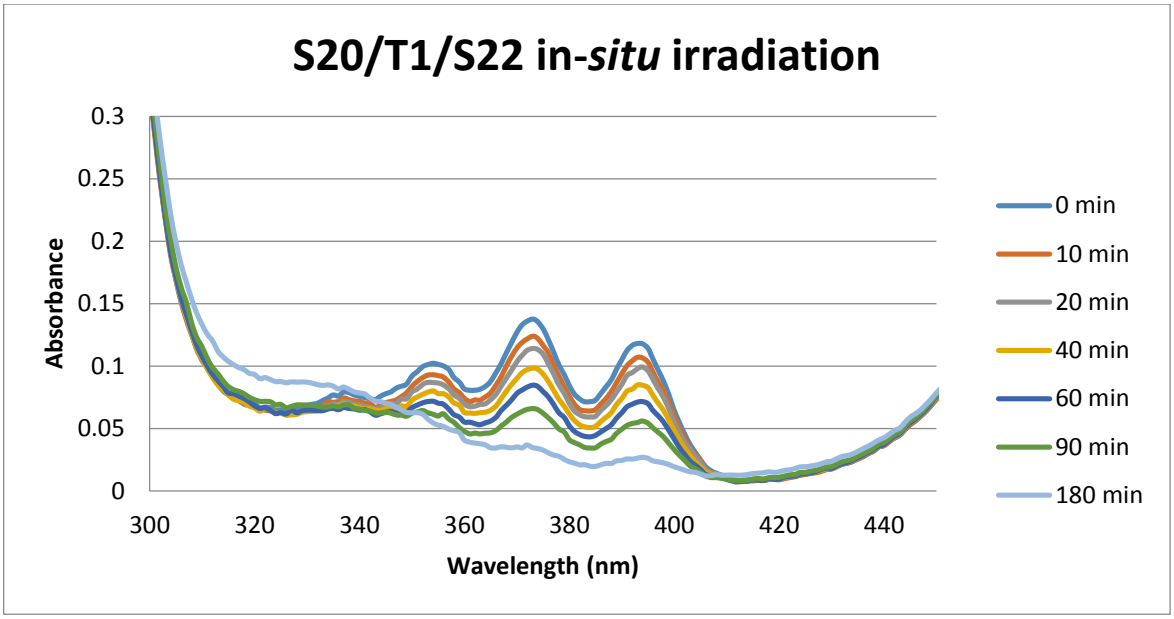


*Figure 17. UV analysis of the in-situ irradiation of **S20**, displaying the successful photodimerisation of the anthracene tags.*

Analysis of the resulting fluorescent data collected throughout the in-*situ* irradiation of these samples did not yield promising results. Figure 18 displays the fluorescence spectra obtained from irradiation of the sample containing the single C/A mismatch sequence **S22**. Despite subtracting the **S22** control data from the scan, Cy5 emission remained in the spectra at the beginning of the reaction, indicating that the strand may already be competing with **S20** and has hybridised with the target, which had been anticipated. However upon photo-irradiation, only a minor increase in the Cy5 emission was observed and the ratio between the Cy3 and Cy5 bands did not change to the extent that would be expected from looking at the control titration in Figure 15. This would suggest that the anthracene photodimer has not weakened the interaction between **S20** and **T1** to an extent that **S22** could compete to fully hybridise with any remaining **T1** that is still bound within the stem-loop. This analysis was again reflected in the data obtained for the in-*situ* irradiation in the presence of the double C/A mismatch sequence **S23**, (Figure 19).

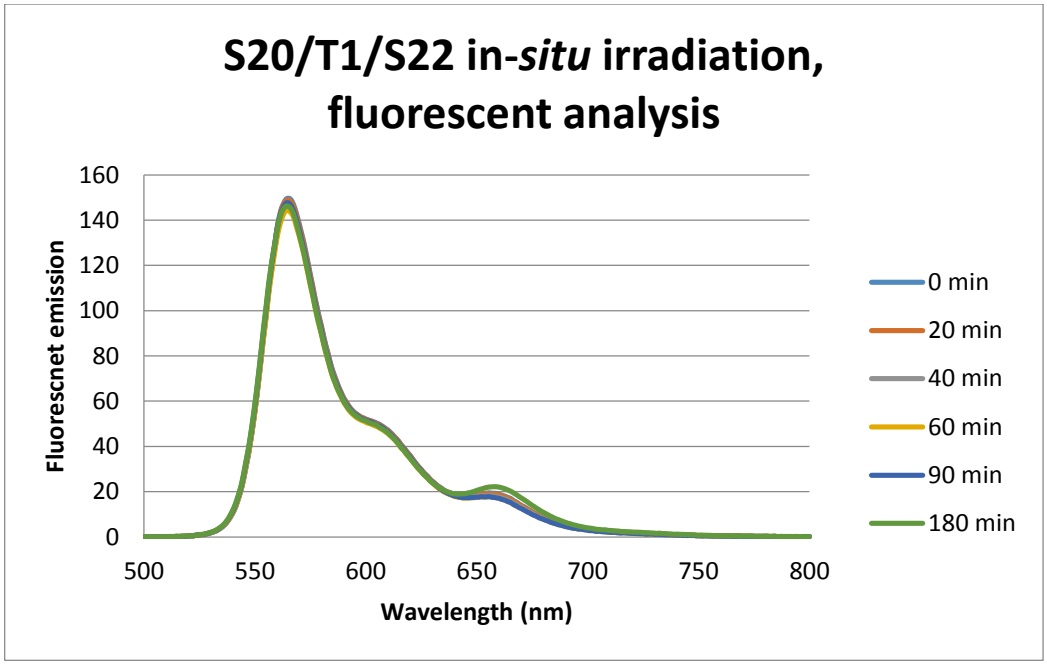


Figure 18. Fluorescence spectra monitoring changes in FRET that are induced through in-situ photodimerisation of anthracene tags in the

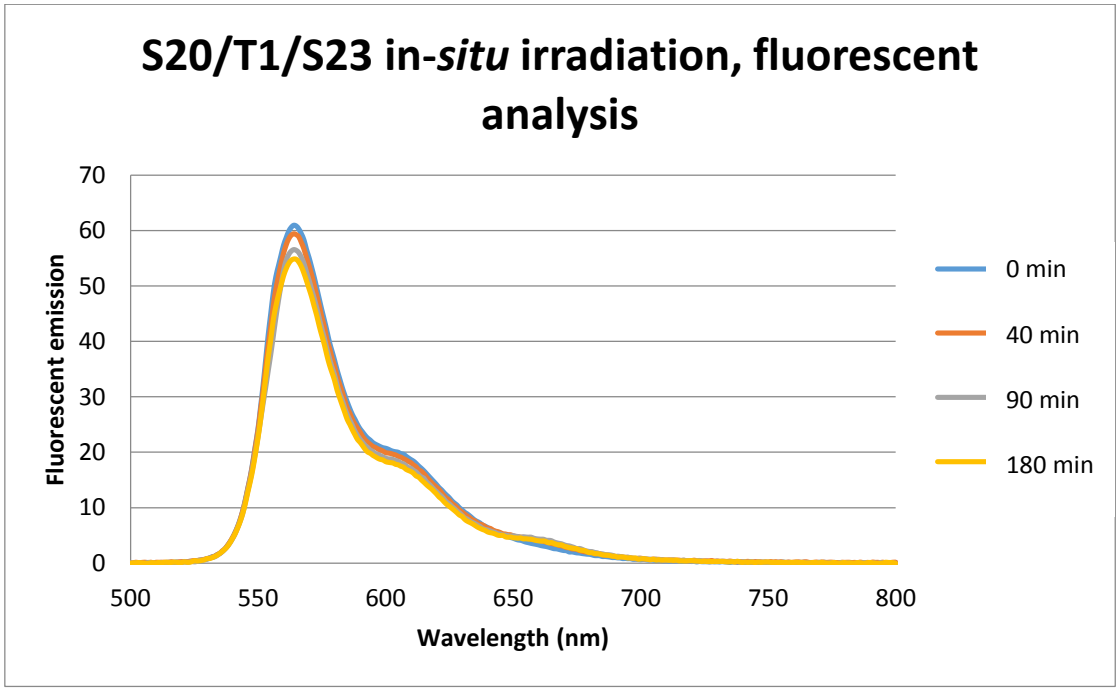


*Figure 19. Fluorescent spectra monitoring changes in FRET that are induced through in-situ photodimerisation of anthracene tags in the presence of **S23**.*

As a further confirmation, the irradiated samples were anaylsed through VT-UV spectroscopy to obtain a melting point curve (Figure 20). In comparison with data

obtained prior to in-*situ* irradiation there did not appear to be any shift in the transition curve that would indicate a switch in the hybridisation of **T1** from the stem-loop to the competing strand and only minor changes in the T_m were observed, (Table 9).

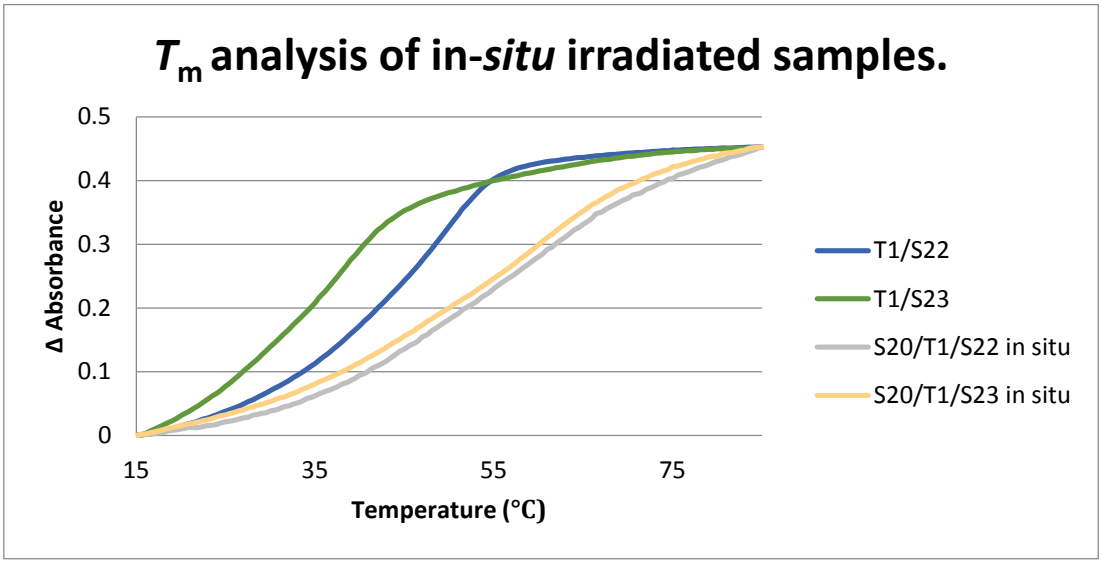


*Figure 20. Melting curve (T_m) analysis of **S20/T1** irradiated in-situ in the presence of the competing strands **S22** and **S23**. Only minor changes in the T_m were observed upon photodimerisation indicating that no sequence displacement had taken place.*

Table 9. T_m data obtained from VT-UV analysis of in-situ irradiated samples.

Sample	T_m (°C)
T1/S22	49.5
T1/S23	39
S20/T1/S22 in *situ*	56
S20/T1/S23 in *situ*	56

4.6 Conclusions and Further Work

The introduction of complementary overhangs between the anthracene and the main oligonucleotide has made advancements on earlier methods to achieve photodimerisation in the presence of the complementary sequence. Whereas attaching two anthracene tags to one strand had previously precluded photodimerisation when in the duplex form (Chapter 3), the presence of overhangs in a stem-loop structure has now

made this possible. However the complex behaviour observed between the stem-loop sequence and the target strand has proved a challenging system with which to obtain 'photo-release'. The stable intra-strand hybridisation between the complementary overhangs, assisted by an interaction with the anthracene tags, hinders the desired opening up of the stem-loop upon addition of the target sequence. Furthermore this is compounded by the target appearing to bind to the stem-loop structure, whether photo-dimerised or not, which prevents a driving force for photo-release. This was an unexpected result when compared with the tried and tested molecular beacon technique, which relies upon the change in secondary structure from stem-loop (no intermolecular duplex) to open form (intermolecular duplex formed).

To further advance this work it would be interesting to investigate different base compositions within the complementary overhangs of the stem-loop sequence. By reducing the number of base pairs, or by lowering the G/C content, the stability of the resulting stem-loop structure could be altered to allow the target sequence to be more competitive. However even if this were successful it is not guaranteed that a shift from duplex to stem-loop upon photodimerisation can fully prevent the target sequence from binding to the free nucleotides within the loop. Therefore additional experimentation may be required to alter the stem-loop sequence by changing the length or introducing certain mismatches. In this way, even if the target could bind to the stem-loop, upon photodimerisation the combination of both the change in structure and mismatches within the loop may be sufficient to completely destabilise the interaction, releasing the target strand.

An additional proposal to further this research is to modify the complementary sequence with additional base pairs so that it overlaps with the self-complementary portion of the stem-loop sequence (Figure 21). In doing so the complementary sequence would be capable of forming additional base pairs, disrupting the stem duplex and potentially driving the equilibrium towards the open state (Scheme 5). However, careful consideration needs to be taken in designing the sequences so that the anthracene overhangs still have enough flexibility to photodimerise once in the fully extended duplex structure. Once again binding of the complementary strand post-irradiation may pose an issue, therefore this strategy could be coupled with the introduction of mismatches, as described previously, in order so successfully achieve photo-release.

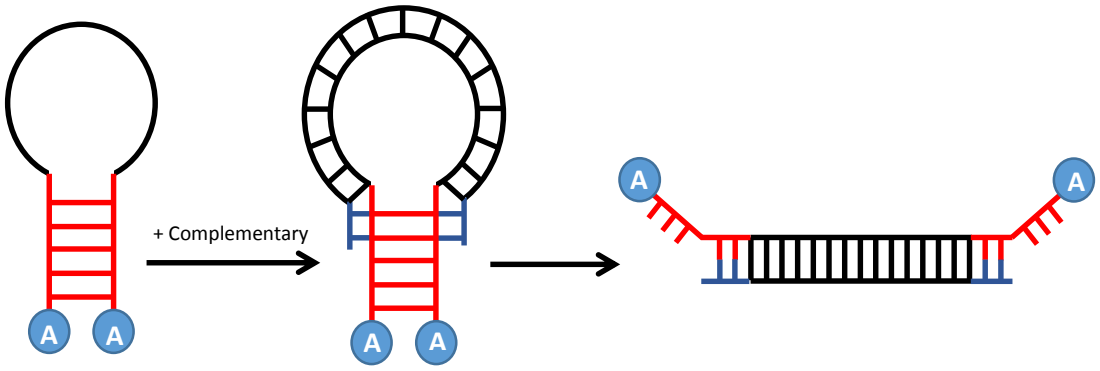
Existing system:

Stem-loop **(S20)**	**X**GCGACTGGACTCTCTCAATGGTCGC**X**
Complementary **(T0)**	ACCTGAGAGAGTTAC

New System:

Stem-loop **(S20)**	**X**GCGACTGGACTCTCTCAATGGTCGC**X**
New Complementary	TGACCTGAGAGAGTTACCA

Figure 21. A new strategy for achieving photo-release through modification of the complementary sequence.



Scheme 5. A representation of how the new complementary sequence competes with the intramolecular base pairs in order to promote opening of the stem-loop structure.

4.7 References

1. Zhang, D. Y.; Seelig, G., *Nature Chem.*, **2011**, *3*, 103-113.
2. Bartlett, J. M. S.; Stirling, D., *Methods in Molecular Biology*, **2003**, *226*, 3-6.
3. Broude, N. E., *Trends in Biotech.*, **2002**, *20*, 249-256.
4. Antao, V. P.; Tinoco, I., *Nucleic Acids Res.*, **1992**, *20*, 819-824.
5. Tyagi, S.; Kramer, F. R., *Nature Biotech.*, **1996**, *14*, 303-308.
6. Lacowicz, J. R., *Principles of Fluorescence Spectroscopy*, Springer Publishing, **2006**.
7. Radi, A.-E.; Acero, J. L.; Baldrich, E.; O'Sullivan, C. K., *J. Am. Chem. Soc.*, **2006**, *128*, 117-124.
8. Li, J. J.; Geyer, R.; Tan, W., *Nucleic Acids Res.*, **2000**, *28*, e52.
9. Whitcombe, D.; Theaker, J.; Guy, S. P.; Brown, T.; Little, T., *Nature Biotechnology*, **1999**, *17*, 804-807
10. Xiao, Y.; Lou, X. H.; Uzawa, T.; Plakos, K. J. I. Plaxco, K. W.; Soh, H. T., *J. Am. Chem. Soc.*, **2009**, *131*, 15311-15316.
11. Fan, C.; Plaxco, K.W.; Heeger, A. J., *Proc. Natl. Acad. Sci.*, **2003**, *100*, 9134-9137.
12. Yu, Z.; Lai, R., *Chem. Commun.*, **2012**, *48*, 10523-10525.
13. Neelakandan, P. P.; Ramaiah, D., *Angew. Chemie. Int. Ed.*, **2008**, *47*, 8407-8411.
14. Blackburn, G. M.; Gait, M. J.; Loakes, D.; Williams, D. M., *Nucleic Acids in Chemistry and Biology*, RSC Publishing, **2006**.
15. SantaLucia, J.; Hicks, D., *Annu. Rev. Biophys. Biomol. Struct.*, **2004**, *33*, 415-440.
16. Integrated DNA Technologies, IDT Biophysics DNA Thermodynamics and Hybridisation tool, Available from: http://biophysics.idtdna.com , [Accessed on 21/01/2014].
17. Yen, W. S.; Blake, R. D., *Biopolymers*, **1980**, *19*, 681-700.
18. Fujimoto, K.; Shimizu, H.; Inouye, M., *J. Org. Chem.*, **2004**, *69*, 3271-3275.
19. Arslan, P.; Jyo, A.; Ihara, T., *Org. Biomol. Chem.*, **2010**, 4843-4848.
20. Zhang, D.; Seelig, G., *Nature Chem.*, **2011**, *3*, 103-113.
21. Clegg, R. M., *Curr. Opin. Chem. Biol.*, **1995**, *6*, 103-110.
22. Stryer, L., *Ann. Rev. Biochem.*, **1978**, *47*, 819-846.
23. http://www.labautopedia.org/mw/index.php/An_Introduction_to_Fluorescence_Resonance_Energy_Transfer_(FRET)_Technology_and_its_Application_in_Bioscience [Accessed on 29/1/14]

Chapter 5: Quadruplex DNA

5.1 Introduction to Quadruplex DNA structures

The guanine nucleobase possesses a unique property in that it can self-associate into cyclic, planar tetramers through Hoogsteen hydrogen bonds in what is known as a G-tetrad. The G-tetrad structure is promoted by the presence of metal cations, which coordinate to the lone pairs of the exocyclic oxygen atoms that point inward towards the centre of the tetramer, (Figure 1A). In the presence of a templating metal cation such as potassium, sequences with a high proportion of guanine can undergo self-assembly to form a four stranded structure known as the G-quadruplex, a stacked arrangement of G-tetrads interlinked by the nucleic acid backbone (Figure 1B).[1] The G-quadruplex structure can occur in a variety of natural and synthetic sequences and can be formed intermolecularly from the association of 2 to 4 individual strands, or more commonly from the folding of a single stranded segment of DNA that loop around the tetrads to generate the final structure.[2]

The structure was first identified in 1962 by Davies and co-workers, but has received renewed attention in recent years due to the discovery of its importance in certain biological roles and in emerging areas such as supramolecular chemistry and nanotechnology.[3,4] The following examples demonstrate how the unique properties of the quadruplex structure can be utilised in a wide range of applications.

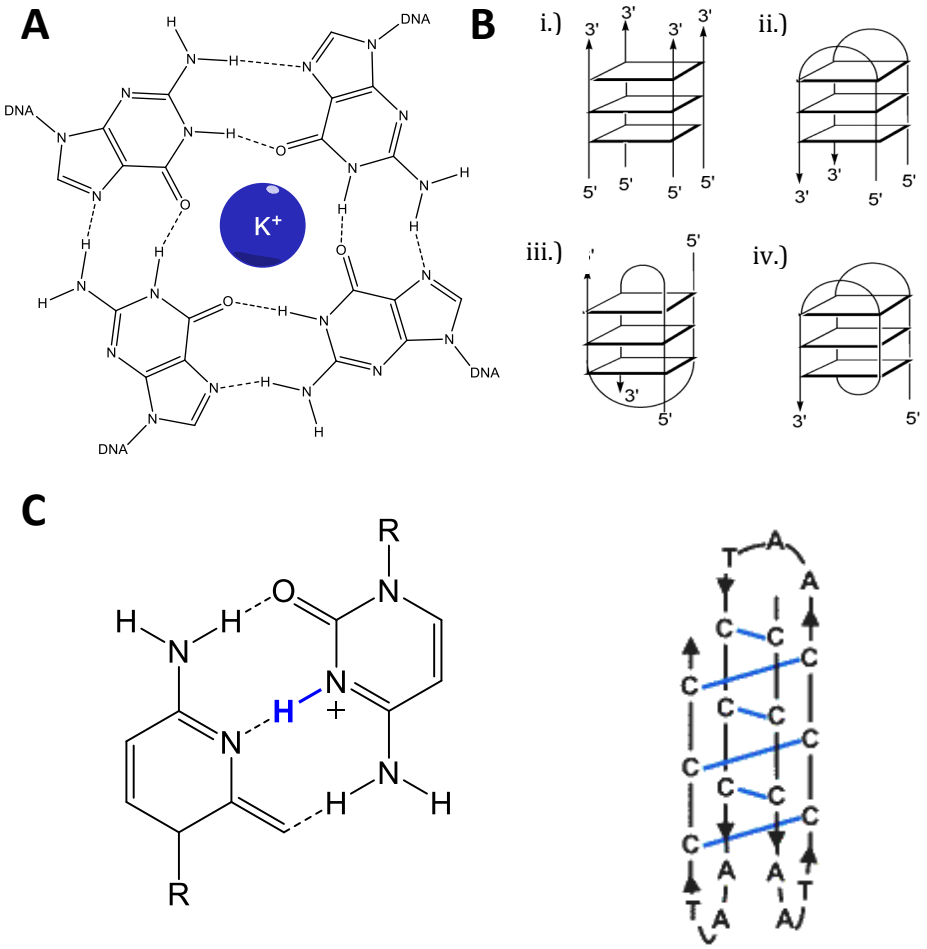
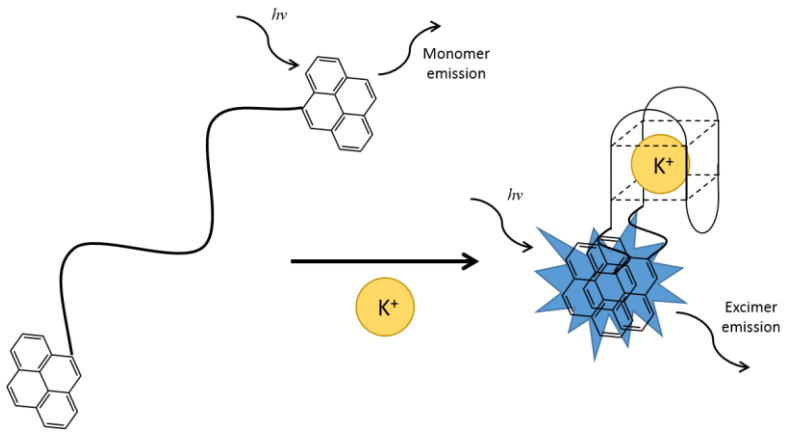


Figure 1. (A) The square arrangement of guanine nucleobases known as a G-tetrad, (B) several G-quadruplex structures including i.) Parallel tetraplex, ii.) Hairpin dimerisation, iii.) Anti-parallel bimolecular complex and iv.) A unimolecular quadruplex. (C) An i-motif structure found in cytosine rich sequences. Folding is dependent on the lowering of the pH so that the cytosine becomes protonated, allowing it to form a C.CH⁺ Hoogsteen hydrogen bond.

The complementary sequence to the G-quadruplex, which contains a high proportion of cytosine is also capable of forming a quadruplex type structure, known as the i-motif. Cytosine can self-associate under lowered pH conditions, protonating the N3 position, allowing it to form a Hoogsteen hydrogen bond to a secondary cytosine. The folded i-motif structure is similar to the G-quadruplex in that it can form from four, two or a single strand, with the cytosine dimers stacking one on top of the other in an alternating fashion to generate a folded structure like that shown in figure 1C.[5]

5.1.1 Fluorescent Sensor for K^+

Due to the conformational changes involved during quadruplex folding, the relative proximity of 3' and 5' termini of an oligonucleotide is altered depending on the type of quadruplex structure involved. Takenaka *et al.* were able to develop a potassium sensing oligonucleotide that uses a FRET based sensing technique to detect the presence of potassium using two fluorescent labels appended to opposite ends of the oligonucleotide.[6] The distance dependant relationship of FRET based fluorescent groups meant that the change in proximity between the termini within the folded structure of the quadruplex resulted in a wavelength shift in the fluorescence spectrum, therefore allowing the presence of the templating cation to be detected (Scheme 1).



Scheme 1. A fluorescent quadruplex based sensor for the detection of K^+ ions based on the excimer fluorescent properties of pyrene.[5]

5.1.2 Nanotechnology

Many of the structural features of the quadruplex have made it an interesting focus of research within nanotechnology, due to its structural compactness and the ability to dynamically control the positioning of molecules on a nanometer scale. Efforts by Tan and Li have sought to use the process of quadruplex folding and

unfolding to create a nanomotor (Figure 2A), capable of performing nano-scale motion through a combination of potassium binding and sequence displacement to drive the relative processes.[7]

Other research within the field has led to the development of DNA ion channels and nano-wires, since short, G-rich strands can be designed to associate into continuous polymers, (Figure 2B).[8] AFM images have shown that G-quadruplex aggregates can grow to 1000 nm in length, where the central core of the G-tetrad acts as a channel through which metal ions can travel, these types of structures therefore have promising applications in nano-electronics and biosensors.

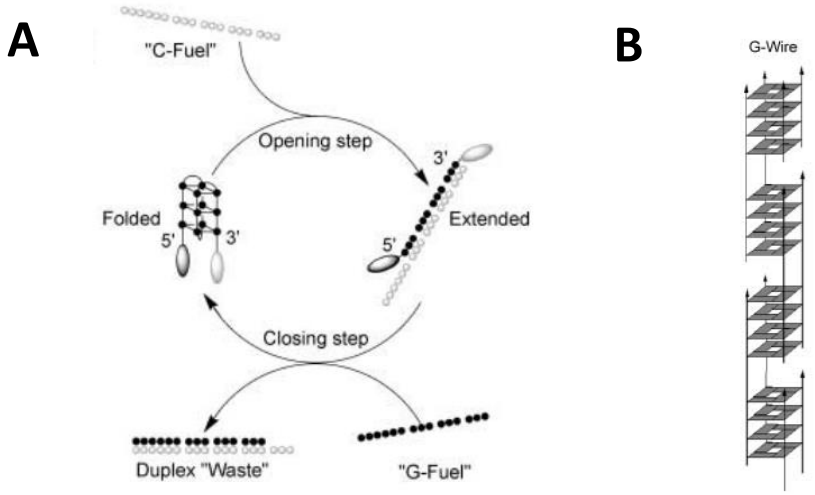


Figure 2. (A) G-quadruplex nanomotor, Image taken from ref. 7 (B) A G-Wire formed from an aggregation of short G-rich oligonucleotides. Image taken from ref. 8

5.1.3 The Telomere and Quadruplex Binding Agents

Regarding their biological role, quadruplex structures have proved to be an interesting target within cancer research. With the advent of genetic sequencing, researchers have noted G-rich segments within the human genome located within single stranded segments at the end of a chromosome known as a telomere.[9,10] In

normal healthy cells the telomere sequence is shortened with each subsequent cell division until such a point that the telomere is so short that the cell undergoes apoptosis and dies. In cancerous cells there is an up-regulation of an enzyme known as telomerase, which can prevent the shortening of the telomere and hence extend the lifetime of the diseased cell.[11]

Since the G-rich sequence within the telomere has the potential to be folded into quadruplex structures researchers investigated whether this structure could inhibit telomerase function and it was found that this was indeed the case. Therefore in an attempt to promote the quadruplex structure within cancerous cells, research is currently underway to develop quadruplex binding ligands that are selective for the G-rich segments within the telomere sequence.[12] There are a number of strategies that have been employed with this target in mind, but the main focus is to bind selectively to a G-quadruplex, increasing its stability and preventing the structure from unfolding in the presence of telomerase. To do this ligand structures are designed to include certain features that maximise the interactions towards quadruplex structures including:

- A large, planar aromatic central structure that can π-stack to the open face of the exterior G-tetrads. Given the large area of the G-quartet quadruplex, ligands typically have a larger aromatic structure than duplex ligands in order to provide selectivity. Further study of these interactions revealed that aromatic structures with a crescent or circular shape maximise overlap and therefore improve π-stacking (e.g. MMQ_3, Figure 3).[13]
- Protonatable pendant arms that will usually include an amine group, which allow the compound to be water soluble as well as being capable of

taking part in electrostatic interactions with the backbone of the quadruplex.[14]

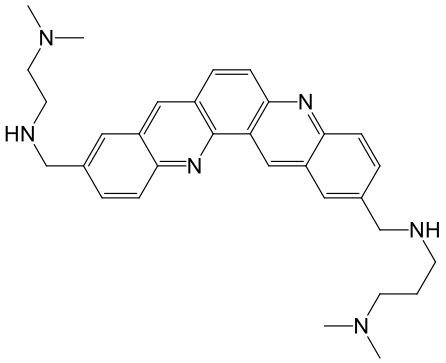


Figure 3. A MMQ_3 ligand designed to target quadruplex structures with a crescent shaped aromatic core and protonatable pendant arms.

As part of this research, many quadruplex ligands have been designed to target and stabilize quadruplex forming promoter regions within the human genome, in particular regions within the genome that have been linked to diseases such as cancer. The formation of a G-quadruplex within genetic DNA prevents the sequence being transcribed by enzymes and therefore the gene is not expressed. The c-*MYC* oncogene is one well known target of this research and the suppression of this gene using a cationic porphyrins such as TMPyP4 has demonstrated effectiveness in limiting the cellular proliferation within cancerous tissues.[15,16]

5.2 Photo-Control of Quadruplex Structures

In addition to the wealth of research being carried out involving quadruplex structures, researchers within the field of nucleic acid chemistry have already begun to investigate the use of photochromic compounds to give an added dimension to G-quadruplex research.

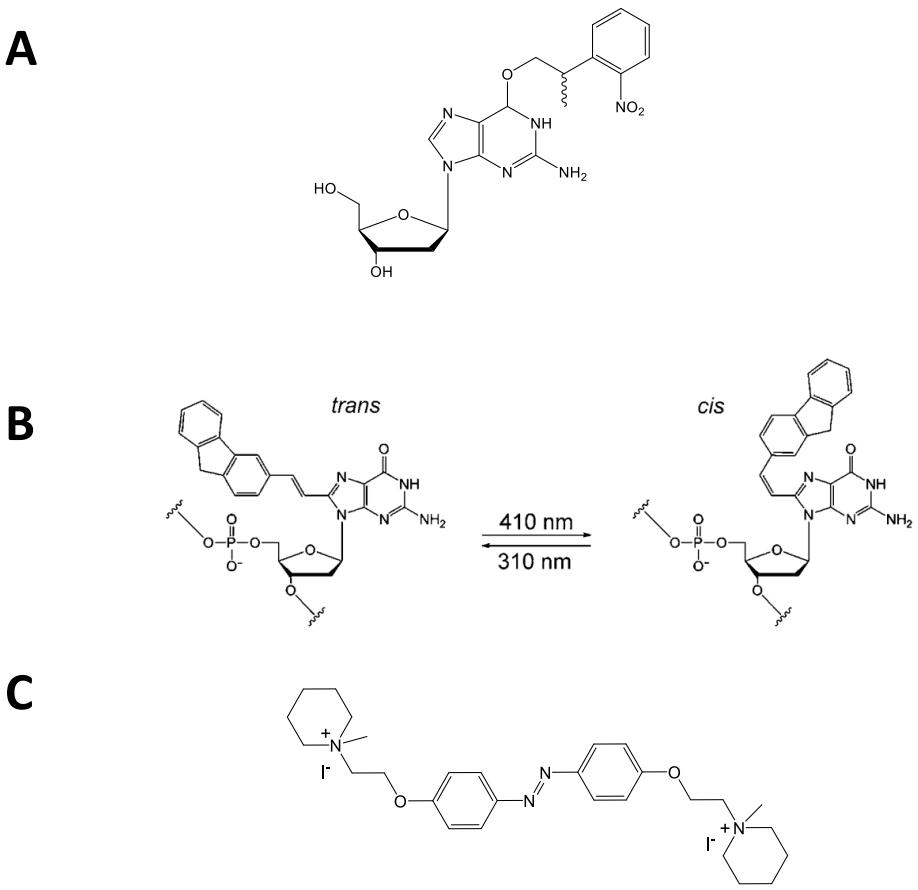


Figure 4. (A) A caged guanine nucleotide, (B) a photochromic nucleobase and (C) a photochromic ligand based on an azobenzene core, all of which can be used in the regulation of quadruplex structures.

Researchers such as Maeda and Heckel, whose earlier work has already been covered in Chapter 1, have attempted to adapt techniques that have already been proven on duplex structures for use on the G-quadruplex. Heckel demonstrated that it was possible to use caged nucleobases placed at certain locations within the sequence to block key interactions such as Hoogsteen type hydrogen bonding and prevent quadruplex folding (Figure 4A). Light triggered release of the caging group therefore allows these interactions to resume, giving temporal control over the formation of the structure.[17,18] Alternatively, Maeda introduced a revised version of the stilbene nucleotide derivative, which provided the reversible photochromic control of quadruplex formation (Figure 4B).[19] G-quadruplex

ligands can also be adapted with photochromic groups as shown by Zhou *et al.*, who functionalised azobenzene with protonatable pendant arms for use as a G-quadruplex promoting ligand, (Figure 4C).[20]

Through cis/trans isomerisation of the azobenzene the binding strength was dramatically altered due to the shift between planar and non-planar structures. This process could greatly effect π-stacking efficiency, and therefore the formation and dissociation of the G-quadruplex could be photo-regulated.

5.3 Thrombin Binding Aptamer.

One of the major issues with G-quadruplex structures is that compared to the standard bimolecular duplex that preferably adopts one type of conformation (A, B & Z), an intramolecular quadruplex can fold into a wide variety of polymorphs, (Figure 5). Depending on a number of factors such as sequence composition, loop structure and the total number of G-tetrads, a singular quadruplex sequence can interconvert between a number of polymorph structures depending on the temperature.[21] G-quadruplex structures are classed as being either parallel or anti-parallel depending on the relative direction of the backbone linking each corner of the G-tetrad. The length and composition of the loop segments that interlink the repeating guanine nucleobases have a major influence on how the different corners of a G-tetrad can orient themselves with respect to each other.

The basket and chair types are examples of anti-parallel (Figure 5A & B), whereas the propeller type is an example of a parallel quadruplex, (Figure 5C). The inter-conversion between quadruplex structures makes it difficult to elucidate the

structure through techniques such as NMR or X-ray crystallography and the sample will often need to be structurally modified or alternatively isolated at low temperatures in order to isolate a single polymorph.[22]

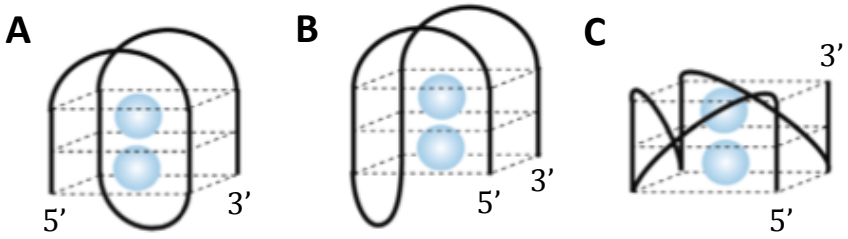


Figure 5. A range of polymorph structures that can be adopted by a G-rich quadruplex sequence including (A) basket, (B) chair and (C) propeller forms.

Modification of quadruplex structures with anthracene or other fluorophores can be a challenging task, as their final position within the quadruplex structure can vary between polymorphs. Therefore the distance between these modifications cannot be guaranteed and any fluorescence and/or photochemical results can often be inaccurate.[23] The quadruplex sequence known as the ‘Thrombin Binding Aptamer’ (TBA) is a short 15 bp sequence with only two G-tetrad stacks (Figure 6).[24,25] It has a distinct advantage over other quadruplex sequences in that it is known to adopt only one polymorph structure, the chair type (Figure 5B). This makes the TBA sequence an ideal candidate for anthracene modification with the objective of controlling quadruplex formation through photodimerisation. The

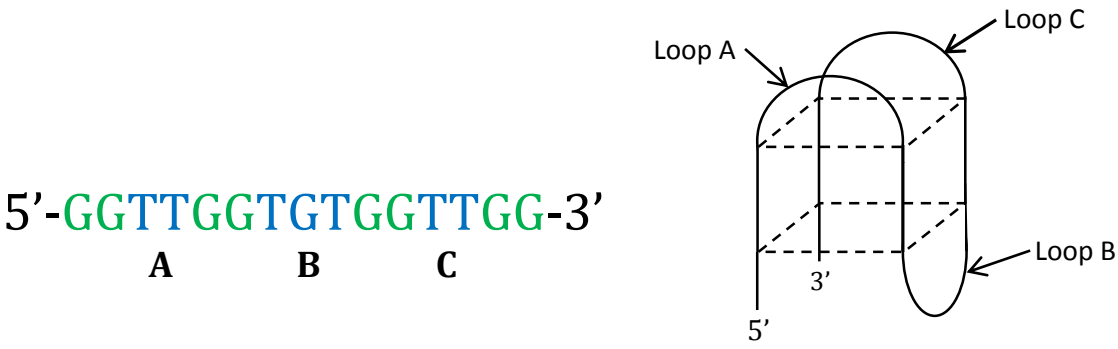


Figure 6. The Thrombin Binding Aptamer (TBA) sequence. Loop segments (A, B & C) in blue and tetrad segments in green.

structure has also been well studied and has been listed on the protein database (1RDE), making it easier to rationalise the position of each anthracene tag within the sequence in order to maximise their proximity once in the folded structure.[26,27]

With careful consideration of the chair conformation and selectively placing anthracene tags in specific locations within the sequence, the relative proximity of the anthracene tags can be altered before and after folding. If this change in distance is significant enough to disrupt photodimerisation, then the reaction could be controlled depending on the presence of a templating metal cation. Alternatively the anthracene tags can be placed so that once photodimerised the sequence will be restricted to such an extent that quadruplex folding is inhibited. An additional feature of the TBA, as implied by the name, is that once in the folded quadruplex form it can bind to certain exosites on the thrombin protein.[28] Thrombin is a serine protease that is an important part of the blood coagulation cascade as it converts soluble fibrinogen into insoluble strands of fibrin, which acts to stop the flow of blood.[29] The binding between thrombin and TBA is mainly achieved through a combination of π-stacking with the external G-tetrad and hydrogen bonding with the free nucleobases within the quadruplex loops, (Figure 7).[30] In attempting to regulate the TBA quadruplex structure through anthracene photodimerisation, it may also be possible to regulate its binding towards thrombin simultaneously. As the binding can only be achieved through cooperative interactions in the folded form, then preventing the TBA from folding through dimerisation should also inhibit binding.

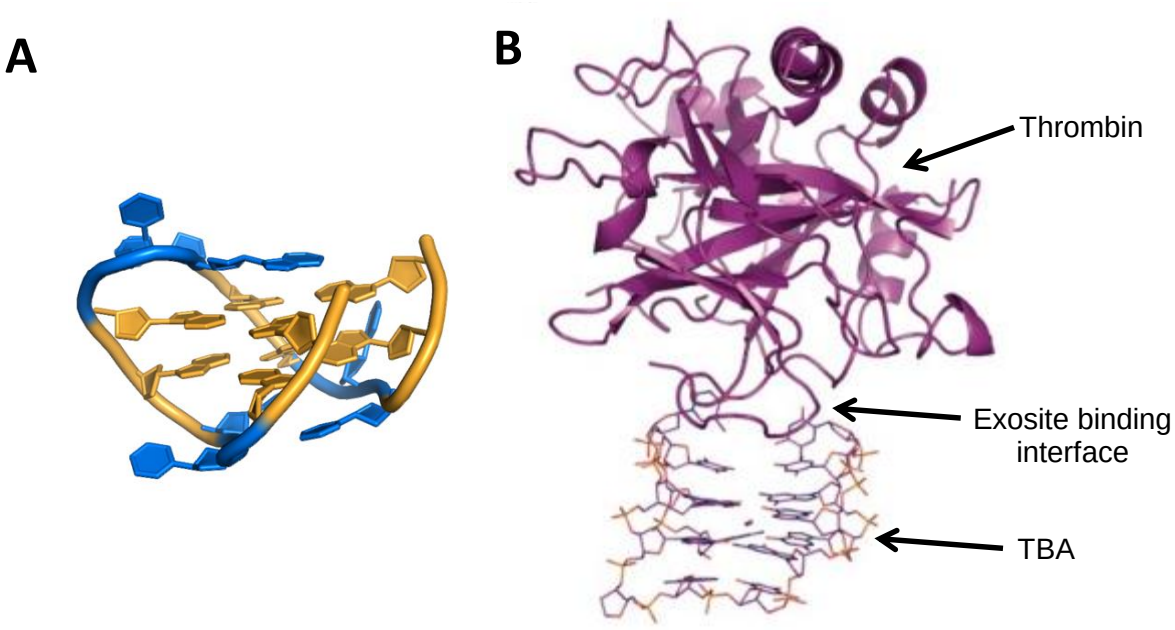


Figure 7. (A) The folded TBA quadruplex containing only two G-tetrads in the chair conformation. (B) Once folded the TBA can bind with certain exosites on the thrombin protein. Image taken from ref. 31.

5.4 Anthracene Modification of the TBA Sequence

In order to utilise the photodimerisation reaction and apply it to controlling quadruplex functionality a number of strategies were devised, which are outlined below. The sequences synthesised to implement these strategies are listed in Table 1.

Table 1. Sequence composition of strands synthesised for TBA quadruplex studies. Sequences have been placed in groups depending on the strategy used and whether the strategy requires K^+ for photodimerisation.

Group	Sequence	Linker Type	Sequence composition (5' – 3')	K^+
1	**S24**	4D	GGTTGGTG**X**GG**X**TGG	N
	S25	4L		
	S26	6D		
	S27	6L		
2	**S28**	4D	GGTTGG**X**G**X**GGTTGG	N
	S29	4L		
	S30	6D		
	S31	6L		
3	**S32**	4D	GGT**X**GGTGTGGT**X**GG	Y
	S33	4L		
	S34	6D		
	S35	6L		
4	**S36**	4D	**X**TGGTTGGTGTGGTTGGT**X**	Y
	S37	4L		
	S38	6D		
	S39	6L		
-	**S40**	-	GGTTGGTGTGGTTGG	-

5.4.1 Strategy one:

In order to inhibit folding of the TBA sequence anthracene tags needed to be placed in proximate positions where they would still be capable of photodimerisation without the presence of a templating metal cation. Due to the fact that the anthracene tags are non-nucleosidic they also needed to be placed within the loop segments of the sequence, leaving the repeating guanine segments intact, which are important for G-tetrad formation. If any of the guanine bases in the sequence were substituted for anthracene the quadruplex may not properly form, therefore the reversible control over the G-quadruplex structure using

anthracene photodimerisation would not be achieved. Two sequences were developed for this strategy, one where anthracene was placed in separate loops B and C (Group 1), and a second where they were placed in the same loop, loop B (Group 2), each utilising a range of anthracene monomers (4D, 6L etc.). Once photodimerised it was hypothesised that the anthracene dimer would restrict these loops so that the correct folding could not be achieved.

5.4.2 Strategy two:

An additional two sequences were created where the anthracene tags are placed in much more distant locations with respect to each other. They were either placed within the outer loops A and C (Group 3), or on the 3' and 5' termini of the sequence (Group 4). In consideration of the TBA's preferred chair-type structure, the anthracene tags would only be brought into close proximity upon folding, and therefore photodimerisation would be dependent upon the presence of the templating cation. Once irradiated the resulting anthracene photodimer should again restrict the flexibility of the oligonucleotide. However in this scenario, the quadruplex would already be in place, and therefore it was predicted that this reduction in flexibility may stabilise the quadruplex structure to some extent and prevent it from unfolding. With sequences in Group 4 the anthracene tag was placed on opposite termini of the sequence; therefore if successfully dimerised it would form a cyclic oligonucleotide, but only in the presence of the templating cation. This cyclic form should in theory not be capable of complete hybridisation in the presence of a complementary strand, as it would be unable to fully extend to form the duplex. This could prove to be useful when applied to the quadruplex

based molecular motor systems discussed in Section 5.1.2 as a method of turning the motor on or off using light.

Sequences in Group 4 were designed with a single thymine extension between the anthracene tag and the outer tetrad guanine nucleobase. This was done to increase the distance between the tag and the lower G-tetrad to minimise possible stacking interactions that could disrupt photodimerisation.

5.5 Results and Discussion.

5.5.1 Photo-Irradiation Studies

To test each of the sequences capability to photodimerise they were subjected to irradiation in the presence or in the absence of potassium, depending on the sequence used. Each sample was monitored through UV spectroscopy to obtain a qualitative idea of the initial rate of reaction and to determine whether any trends could be observed in the data that would suggest an optimal linker configuration.

The first observation that was made was that out of all the sequence groups within this study, only sequences within Group 3 displayed no signs of reactivity, with little to no change in the anthracene absorption band over time when irradiated in the presence of potassium. It was clear that the loop structures within the TBA quadruplex are too distant for the anthracenes to reach in order to photodimerise, or they are oriented in such a way that in all cases the tags are directed away from each other. Therefore no photodimer was collected for these sequences and they were not tested further.

For the other TBA sequences, which did display reasonable reactivity, the rate at which they photodimerise was again dependent on how they were placed within the sequence. Sequences within Groups 1 and 2, where the anthracene is placed either one or two bases apart, gave the best results (irradiated in the absence of potassium); whereas sequences in Group 4 that are dependent on the folding of the quadruplex to increase proximity (irradiated in the presence of potassium), were slightly slower to react (Figure 8A-C). Using the 1RDE pdb file of the folded TBA quadruplex the distance between the 3' and 5' termini was measured to be ~18 Å. Compared with the respective 10.2 Å and 6.8 Å separation between tags in Group 1 and 2 sequences, this measurement could explain the trend observed in the data, although some other factors could come into play such as disruptive π-stacking interactions.

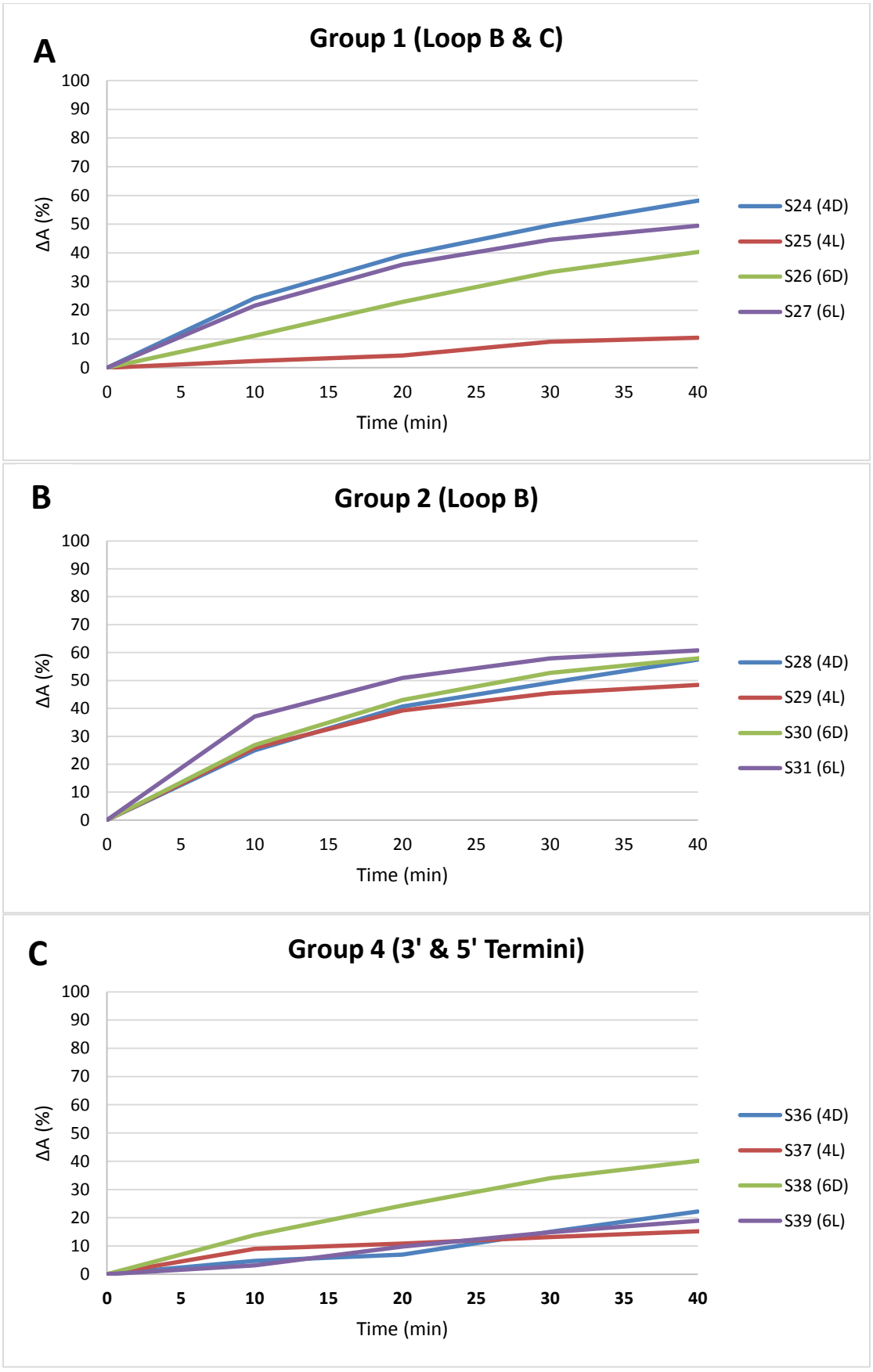


Figure 8 (A-C). Changes in the absorbance (365 nm) measured upon irradiation of anthracene modified sequences, giving an indication of sample reactivity.

For Groups 1 and 2 the change in linker type did not have an overwhelming effect on the reactivity, with most sequences displaying relatively similar changes in absorption over time. The one exception being **S25**, where the 4L linker proved to hinder the photodimerisation reaction, (Figure 8A). As for Group 4, there was a substantial improvement in reactivity when using the 6D linker, where the additional length and better orientation appeared to provide an increased chance of collision between the anthracene tags (Figure 8C).

In order to show that the folding of the TBA quadruplex had an effect on the rate of photodimerisation, sequences **S36** and **S38** in Group 4 were tested with and without the presence of the templating potassium cation. The resulting data (Figure 9) displays a significant decrease in the reactivity in the absence of potassium for both sequences. The cation induced folding of the quadruplex has therefore substantially improved the proximity between anthracene tags compared to the random coil state of single stranded DNA.

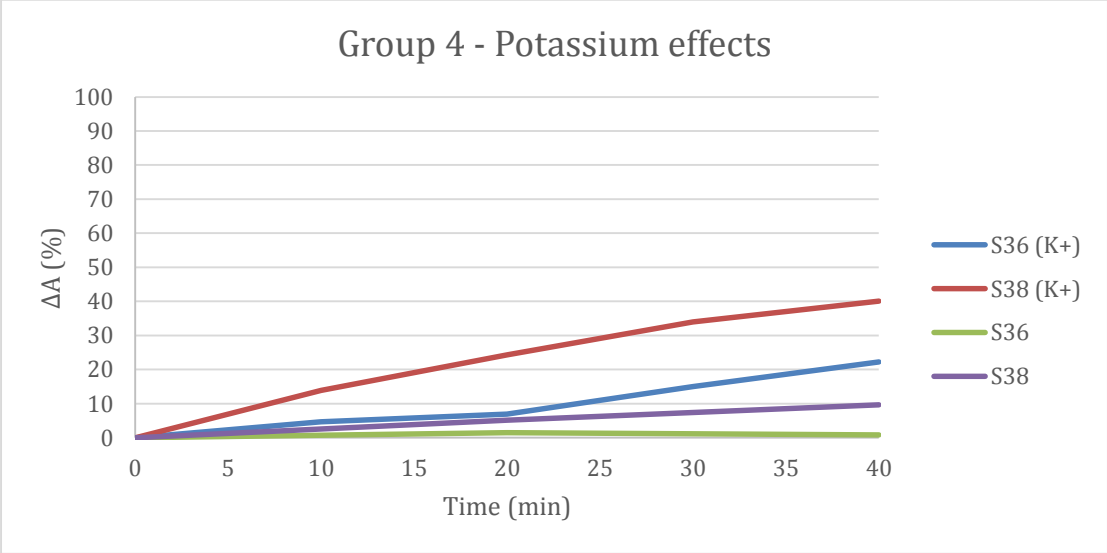


*Figure 9. An irradiation profile measuring the effects of potassium on the rate of reaction of group 4 sequences **S36** and **S38**. This data suggests that the binding of K^+ increases reactivity due to the folding of the G-quadruplex.*

5.5.2 HPLC Results

After irradiation, each sample was analysed by analytical HPLC and the resulting data supports what was observed from the irradiation studies, where sequences with improved reactivity had higher yields of the D_1 and D_2 photoproducts (Table 2). As seen in earlier studies (Chapter 3.4.1), the ratio between the two types of photodimer depends mainly on the forward rate of reaction; for samples where the reaction proceeded to the products more readily, there was a higher ratio of the more thermally stable D_1 photodimer. One interesting exception to this result is for sequence **S37** (Group 4) where the photodimerisation produced a single photoproduct by HPLC. It is already known that the gap separating the two tags is much greater for these folded sequences, but the use of the 4L linker within this particular sequence may orient the tags in such a way that the formation of the less thermally stable D_2 (Head to Head) photodimer is inhibited.

Table 2. Relative area of peaks observed in HPLC analysis immediately after irradiation (40 mins).

Group	**Sequence**	**HPLC Relative Area (%)**			
		Starting material	D_1 (HT)	D_2 (HH)	Ratio (D_1/D_2)
1	S24 (4D)	46.5	41.4	12.1	3.42
	S25 (4L)	54.7	33.0	12.3	2.68
	S26 (6D)	49.0	39.0	12.0	3.25
	S27 (6L)	26.2	65.5	8.3	7.89
2	S28 (4D)	53.5	22.2	24.3	0.91
	S29 (4L)	57.9	8.7	33.4	0.26
	S30 (6D)	40.2	37.4	22.4	1.68
	S31 (6L)	73.4	10.9	15.7	0.69
4	S36 (4D)	72.4	10.3	17.3	0.60
	S37 (4L)	89.1	10.9	-	-
	S38 (6D)	58.1	15.5	26.4	0.59
	S39 (6L)	75.3	11.4	13.3	0.86

The linker length appears to be play a role in determining the reactivity of a sequence, with longer linkers providing the additional flexibility required to initiate slightly more efficient photodimerisation; within the three groups, sequences that utilise the n = 6 linker provided the highest yields of the photodimer. The configuration also appeared to play a role for the longer n = 6 linkers, with the D isomer reacting more readily for Group 2 and 3 but less readily for Group 1. In addition the D_1/D_2 ratio was affected to some extent by the configuration used.

5.5.3 Melting Point (T_m) Analysis

The thermal denaturation of quadruplex structures was analysed in the same way as the duplex structures, where a change in absorbance was observed as the temperature is increased and the quadruplex unfolds. However, for quadruplex structures, there is a decrease in absorbance measured at 297 nm as the temperature is raised and therefore a hypochromic transition is observed in the VT graph (Figure 10).[32-34] Despite this difference it can still be used to determine the T_m value, which for these structures, defines the temperature at which 50% of DNA is folded and 50% is unfolded. Each sample was analysed by T_m analysis before and after photodimerisation in the presence of potassium and compared to obtain a ΔT_m value (Table 3). The photodimer for each sample was purified by HPLC before analysis.

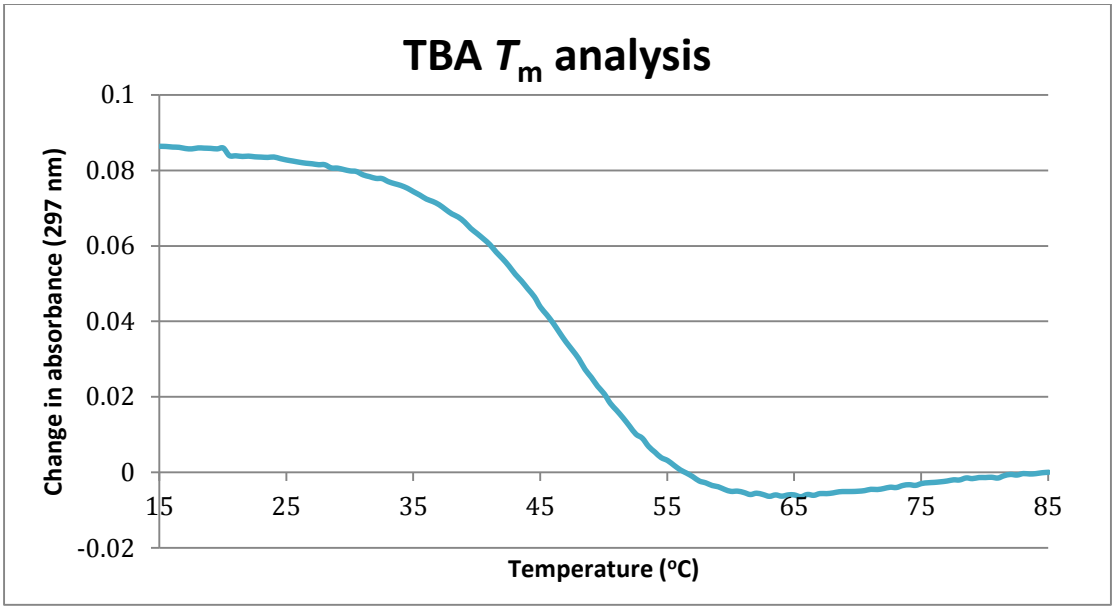


*Figure 10. T_m analysis of a TBA sample **S40** (5 µM sample, 50 mM KCl, 10 mM Tris Buffer). The sigmoidal transition displayed is hypochromic, a decrease in absorbance at 297 nm when the temperature is raised.*
***S40**:* GGTTGGTGTGGTTGG

Un-photodimerised structures:

The first major observation taken from the T_m data is the comparison of each sample to the unmodified sequence (**S40**) prior to irradiation, which can be used to determine the effect anthracene modification has had on the quadruplex structure. For sequences in Group 1 and some of those in Group 2, the modified sequences have a similar T_m to **S40** and this observation can easily be explained. Unlike the modification of duplex structures, none of the bases that are substituted for an anthracene tag are involved in the principle hydrogen bonding interactions that form the quadruplex secondary structure. Therefore any variation in T_m between the samples and **S40** will mainly depend on where the anthracene is located within the final folded structure and whether they can take part in minor interactions that can stabilise or destabilise the quadruplex.

*Table 3. T_m data collected for TBA sequences **S24** to **S40** (5 μM oligo concentration, 50 mM KCl, 10 mM Tris Buffer, 0.5 °C/min ramp speed). No T_m transition was observed for D_1 the photodimers within group 1.*

Group	Sequence name	Sequence	Linker	State	T_m (°C)	ΔT_m (D_1-SM)
1	S24	GGTTGGTG**X**GG**X**TGG	4D	SM	37 (±1.4)	-
				D_1	-	
	S25		4L	SM	37.5 (±1.4)	-
				D_1	-	
	S26		6D	SM	46.5 (±1.0)	-
				D_1	-	
	S27		6L	SM	47.5 (±0.7)	-
				D_1	-	
2	S28	GGTTGG**X**G**X**GGTTGG	4D	SM	45.5 (±2.1)	-0.5
				D_1	45 (±1.4)	
	S29		4L	SM	46 (±2.1)	-1.0
				D_1	45 (±1.4)	
	S30		6D	SM	70 (±1.4)	-12
				D_1	58 (±1.4)	
	S31		6L	SM	49.5 (±0.7)	-11.5
				D_1	38 (±1.4)	
4	S36	**X**TGGTTGGTGTGGTTGGT**X**	4D	SM	52 (±1.4)	-2.5
				D_1	49.5 (±0.7)	
	S37		4L	SM	54 (±2.1)	-4.5
				D_1	49.5 (±0.7)	
	S38		6D	SM	62 (±2.1)	-6
				D_1	56 (±1.4)	
	S39		6L	SM	62 (±1.4)	-6.5
				D_1	55.5(±0.7)	
-	S40	GGTTGGTGTGGTTGG	-	-	44.5 (±0.7)	-

Sequences within Group 4 and sequences **S30** and **S31** from Group 2 gave a significant increase in the T_m compared with the unmodified sequence (**S40**). This increase in T_m ranges from 5 °C to as much as 25 °C and can be justified by significant π-stacking interactions between the anthracene tag and the external face of the G-tetrad stack. In comparison to Group 1, sequences within Groups 2 and 4 have the anthracene positioned so that they stack to the same G-tetrad of the TBA quadruplex. Therefore the substantial stabilisation to the quadruplex

observed for these sequences could be a result of cooperative π-stacking interactions from both anthracene tags.

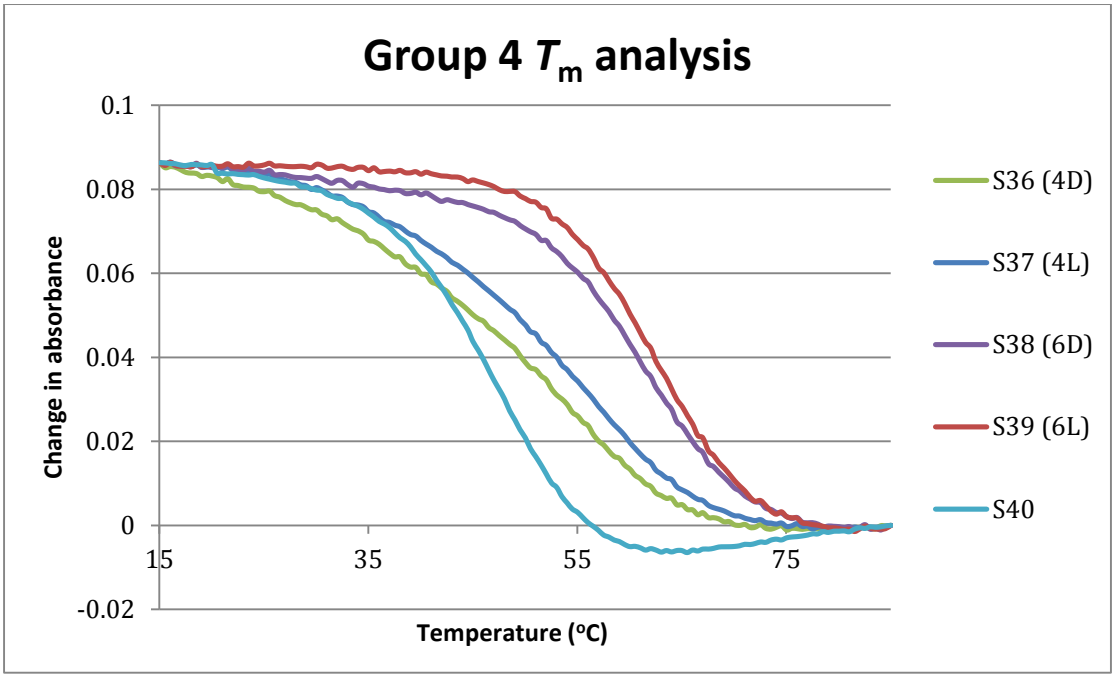


*Figure 11. T_m data for group 4 sequences displaying variations between the n = 4 and n = 6 linkers. It is clear that n = 6 linkers are better at stabilising the quadruplex. Overall there is an improvement in the T_m compared with the unmodified sequence **S40**.*

The data suggests that the extent of the π-stacking interactions is affected by the type of linker used for each sequence. Sequences in Group 4 that utilise the *n* = 6 linker display T_m values of around 10 – 25 °C greater than that of their *n* = 4 counterparts, (Figure 11). Presumably the additional length provided by the *n* = 6 linker allows the anthracene tag to adopt the most favourable position to overlap with the G-tetrads, maximising their interaction with each other. In the case of sequence **S30** the 6D anthracene tags lead to a T_m of 70 °C and therefore the arrangement of the anthracene tags within this sequence leads to the greatest stabilising interactions with the G-tetrad. This observation was subsequently

highlighted in CD analysis (Section 5.6 – Figure 15), which displayed additional bands within the spectrum observed solely with sequence **S30**.

For all of the sequences analysed within this study the stereogenic configuration of the anthracene linker does not have a significant impact on the resulting T_m data. The only exception to this observation is sequences **S30** and **S31** within Group 2 where there was a 21 °C difference in stability for the D isomer over the L. This difference was apparent for both the starting material and photodimer samples and subsequently the change in T_m (Δ) upon photodimerisation for both samples was comparable. It is therefore clear that this difference in T_m is not due to anthracene π-stacking interactions and is most likely a combination of factors when using the 6D and 6L linkers in conjunction with this sequence, as the same results were not observed when using the n = 4 linker (**S28** and **S29**). The data observed for these sequences could be rationalised by considering which groups with the anthracene modification interact with the quadruplex structure. The longer n = 6 alkyl spacer is bulkier and could be disruptive to the quadruplex if positioned in a conflicting orientation, which would explain why the T_m of the L sample (**S31**) drops below the control (**S40**) after dimerisation. Conversely the D sample (**S32**) would be oriented in the opposite orientation and would not display these interactions, giving it greater stability over the L sample.

Photodimerised structures:

With regard to the strategies set out in Section 5.4 to control the quadruplex structure using anthracene photodimerisation there were a number of trends that were expected from the T_m data. Firstly a net reduction in the T_m value for sequences in Groups 1 and 2 was expected as the dimerisation inhibits folding and destabilises the quadruplex structure. Secondly those sequences within Group 4 were expected to have T_m values that either remain unchanged or even increase due to the covalent linking of the ends of the oligonucleotide, which could potentially prevent unfolding of the quadruplex.

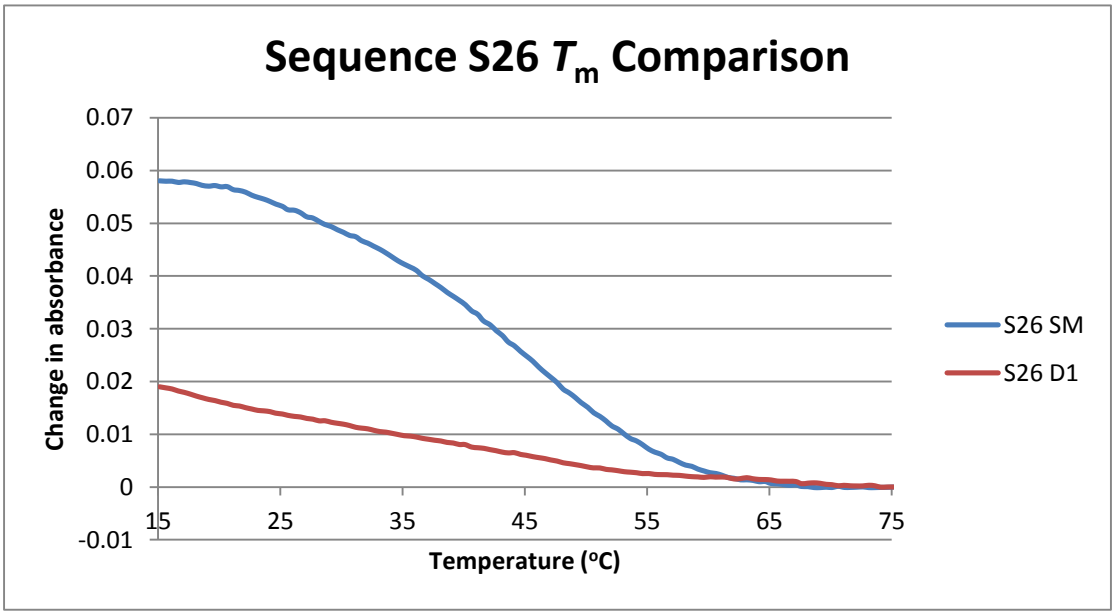


*Figure 12. T_m data for sequence **S26**. Upon photodimerisation the quadruplex folding transition is no longer observed.*

***S26**:* GGTTGGTGXGGXTGG

***S26-D1**:* GGTTGGTGXGGXTGG

The data in Table 3 shows that whilst some of these predictions were observed, some other trends were also apparent. Sequences within Group 1 displayed a

dramatic change upon photodimerisation in that no quadruplex folding/unfolding transition could be observed when varying the temperature (Figure 12). This is a good indication that photodimerisation of anthracene tags between separate loops has had the desired effect of restricting the orientation of the oligonucleotide, preventing it from adopting the quadruplex structure in the presence of potassium. Rather than reducing the overall stability of the quadruplex structure the photodimerisation has in effect switched off quadruplex folding (Figure 12), making the Group 1 sequences good candidates for further research in a number of applications.

As for sequences from Group 2, the degree of change to the T_m value (Δ) that was observed depends greatly on which type of linker is used for the particular sequence. Those which utilise the shorter n = 4 linker display only a minor change in T_m, and this varies little from the unmodified sequence **S40**. This indicates that the anthracene has little impact on the quadruplex structure and the dimer does not place any significant restriction on the flexibility of the oligonucleotide, allowing it to fold and unfold freely. However for sequences that utilise the longer n = 6 linker there are more substantial changes to the T_m upon photodimerisation, although the explanation for this can be attributed to the interactions between the anthracene tag and the G-tetrad rather than any restriction in flexibility.

5.6 Circular Dichroism Studies.

The chair-type conformation of the TBA aptamer can readily be identified through CD spectroscopy by the appearance of three characteristic bands, two positive at 295 and 245 nm as well as a negative band at 270 nm, depending on the

anthracene modification.[35] Titration of a TBA sequence with the templating potassium cation induces the folding and stabilises the quadruplex structure, increasing the intensity of these bands (Figure 13). The CD profiles have isoelliptic points at 280 and 255 nm, indicating the two-state nature of the structural transition of the quadruplex upon potassium ion binding.[36]

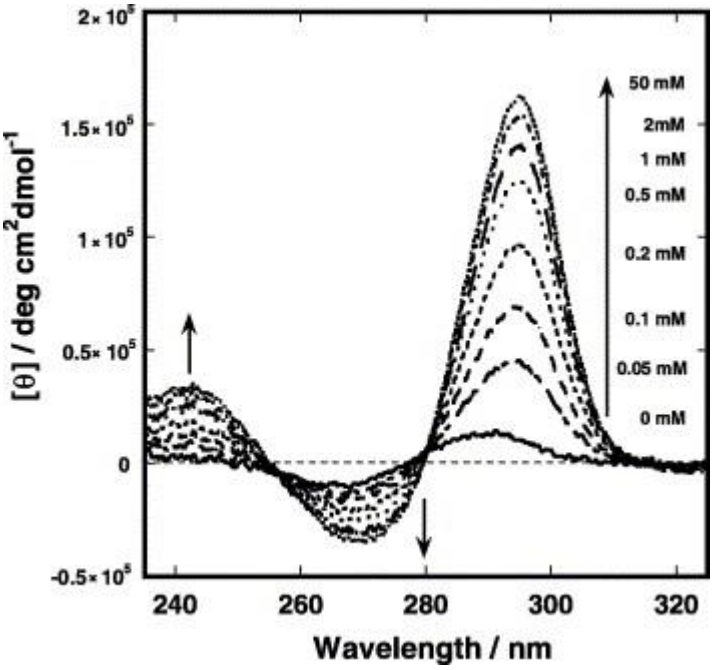


Figure 13. A CD titration of a TBA sequence with K^+, which induces quadruplex folding. The quadruplex can be identified by three characteristic bands at 245, 270 and 295 nm. Image taken from reference 37.

As previously stated in observations from T_m analysis, the TBA quadruplex structure can be further stabilised by interactions between the anthracene tags and the external face of the G-tetrad. This interaction can be observed through a shoulder peak at 265 nm and a shift in the absorbance maxima wavelength of the negative CD band from 270 nm to 250-260 nm. For most sequences the shoulder was not very pronounced and was only observed as a slight sharpening of the negative band (Figure 14). However for sequence **S30,** where there was a significant interaction leading to a 25.5 °C stabilisation over the unmodified

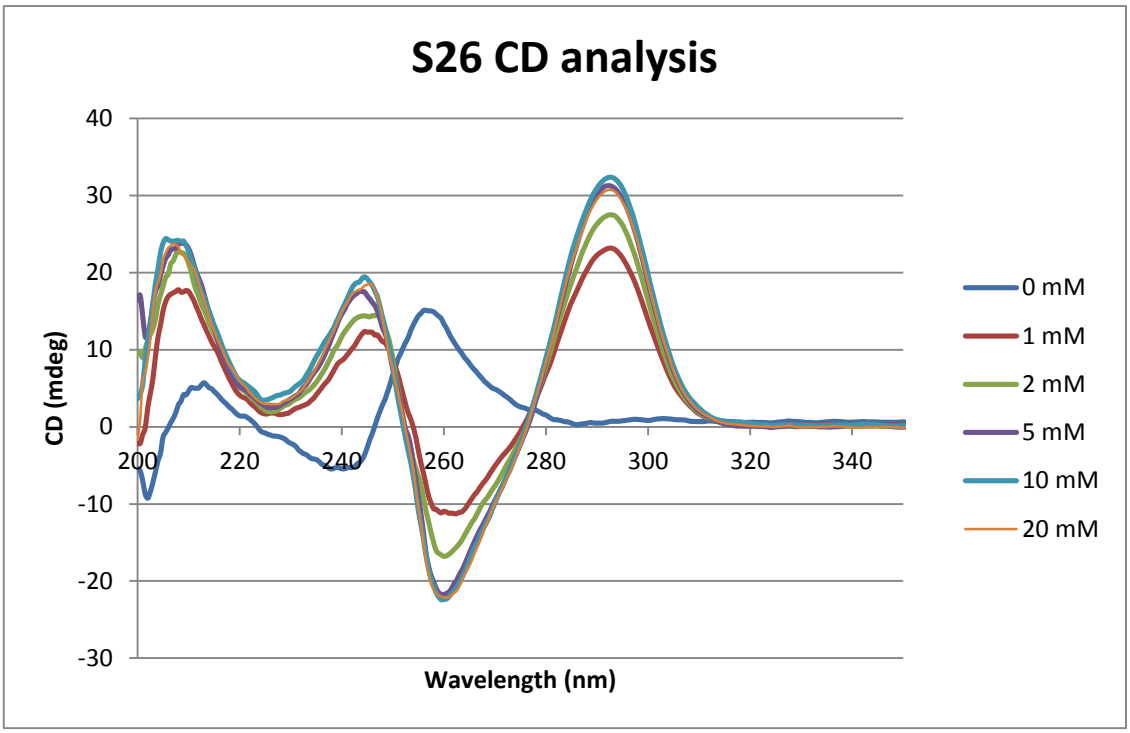


*Figure 14. CD titration of an anthracene modified TBA sequence (**S26**) against K^+. The quadruplex folding remains unaffected upon addition of K^+ and an induced anthracene signal can be seen at 260 nm observed as a sharpening of the negative band.*

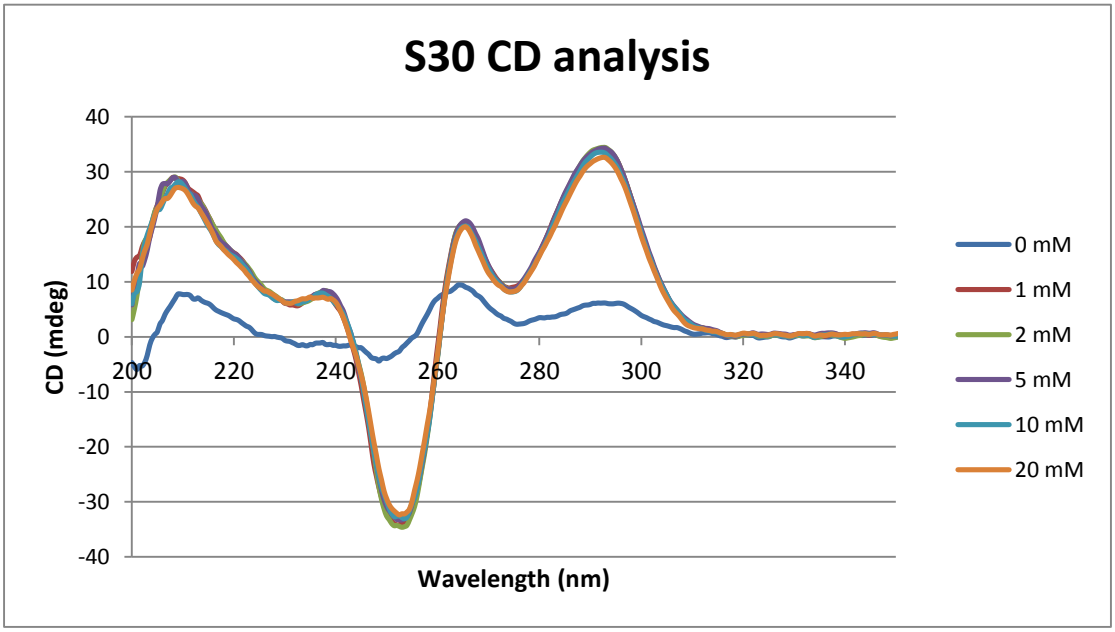


*Figure 15. CD titration of **S30** against K^+, a pronounced anthracene band can be observed at 265 nm, which occurs only with this sequence.*

sequence, there was a significant shoulder peak observed within the CD spectrum (Figure 15). It is possible that this is due to a cooperative interaction using both

anthracene tags, which is only possible with the 6D variant of the threoninol linker.

On analysis of the photodimer quadruplex samples, it was apparent that the observed CD data correlated with results obtained earlier from T_m experiments. Sequences within Group 2 did not show any variation in T_m upon photodimerisation and therefore displayed the typical transition into a quadruplex structure upon addition of the templating potassium ion despite the photodimer being present. However sequences within Group 1, which did not exhibit any T_m transition upon photodimerisation displayed a markedly different behaviour under the same conditions.

An initial observation of particular interest is that the initial CD spectrum analysed before the addition of potassium displayed different band patterns before and after irradiation, (Figure 16). Whilst the starting material displays bands that are characteristic of random coil DNA, after dimerisation the bands shift into what appeared to be the three banded signal of the TBA quadruplex. However upon the subsequent addition of the potassium ion there were only minor changes in the positive band at 295 nm and not the expected transitions representative of quadruplex folding, (Figure 17). Therefore while this result may confirm that quadruplex folding is inhibited by the anthracene photodimerisation, the data suggests that the same process has initiated a change in the structure of the single stranded sample similar to the folding that would occur when adding potassium to the sample. However these two types of folding do not appear to be the same and the anthracene photodimer inhibits any further significant change when

potassium was added to the sample. This data is a good indication that the photoproducts of sequences from Group 1 are restricted from quadruplex folding by the anthracene dimer, which correlates with the T_m data.

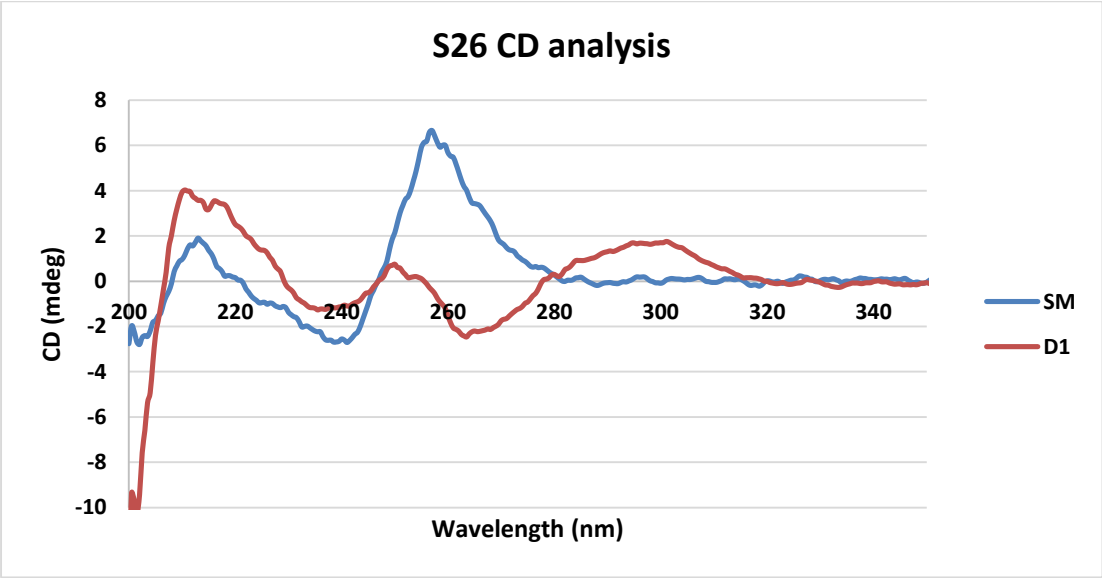


*Figure 16. CD analysis of **S26** before and after photodimerisation and in the absence of a templating cation (K^+). The data displays a shift in the bands from normal random coil upon photodimerisation.*

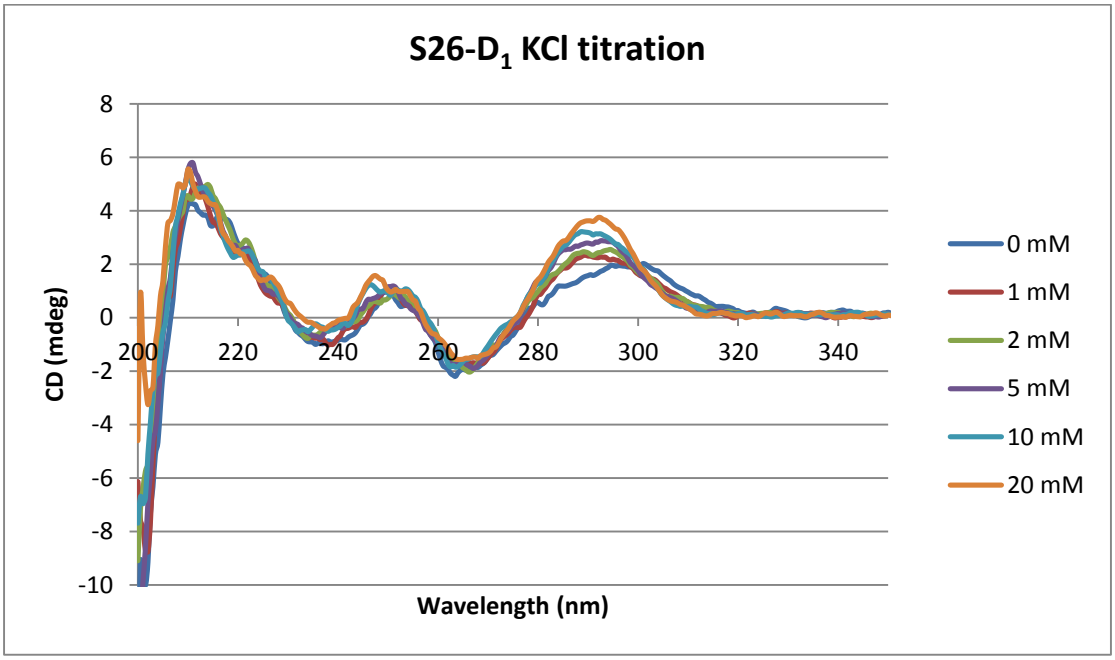


*Figure 17. A CD titration of the photoproduct **S26-D1** against K^+. The data displays only small shifts in the positive band from 300 to 290 nm and not the changes representative of quadruplex folding.*

On analysis of other photodimerised sequences from group 2, which displayed less significant changes in T_m, normal G-quadruplex folding was observed upon addition of potassium as seen in the CD titration, (Figure 18). This confirms the T_m that the anthracene is not capable of inhibiting folding using this strategy of placing the tags within the same loop with a single base separation.

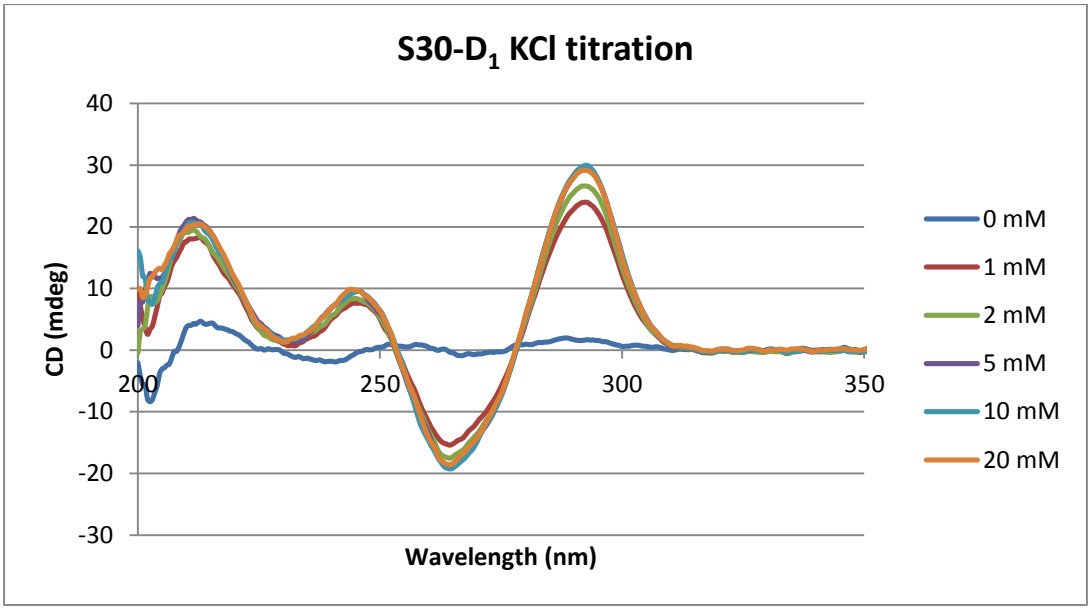


*Figure 18. CD titration of **S30-D1** from group 2. Normal G-quadruplex folding occurs upon addition of potassium and resembles the initial titration displayed in figure 15, with the exception that it no longer displays an anthracene shoulder peak now that it is in the form of a photodimer.*

5.7 Thermal reversion

To test the reversibility of the anthracene photodimer within the TBA sequences, samples were analysed at an elevated temperature (80 °C) using a VT-UV spectrometer to measure the increase in absorbance as the anthracene photodimer dissociated (Figure 19). The procedure summarised in Chapter 3.7 evaluates the thermal stability of the photodimer by comparing the rate of reversion (k_{diss}), which can be calculated from the reversion data using the formula outlined in Chapter 3, Equation 1. Samples **S26** and **S38** from Groups 1

and 4 respectively were analysed using this technique and the dissociation rate constants for each sample are listed in Table 4. Both samples were reasonably resistant to thermal reversion, requiring overnight heating (approximately 18 to 20 hours) to achieve full reversion to the starting material, which was confirmed through analytical HPLC. These results display a faster rate of reversion than those observed previously involving samples designed to manipulate duplex binding (Chapter 3.7), this indicates the possibility of more structural strain placed on the photodimer within these samples. This is particularly apparent with sequence **S38**, where the photodimer joins opposite ends of the sequence to form a cyclic oligonucleotide. Under increased temperatures, the influence from the oligonucleotide as it attempts to unwind from the cyclic form serves to increase the rate of dissociation, giving the highest value for k_{diss} seen so far of 3.8 s^{-1}.

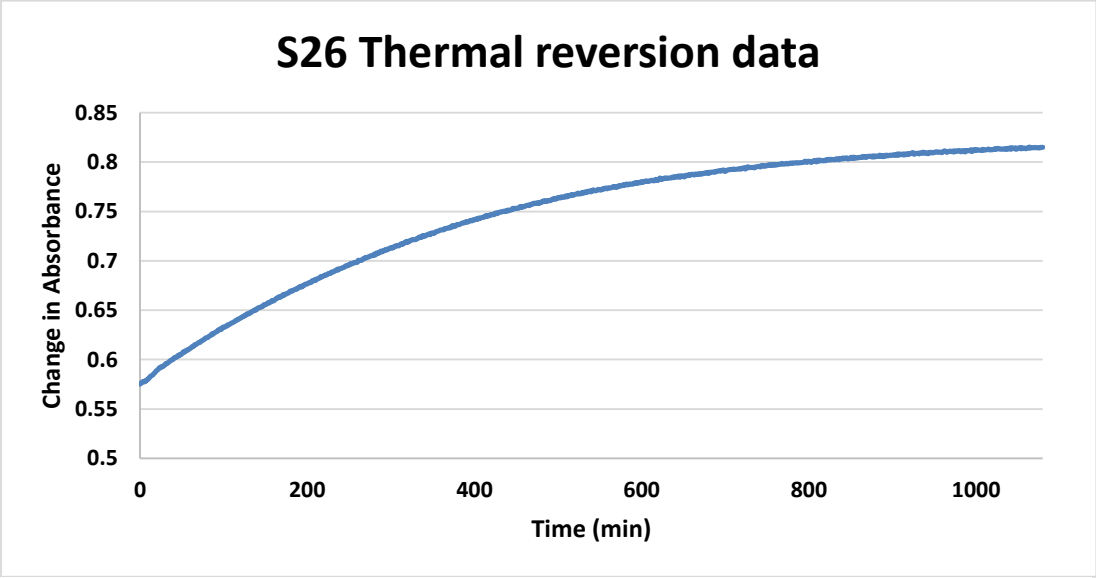


*Figure 19. Thermal reversion data of sequence **S26** (80 °C).*

Table 4. Rates of reversion (k_{diss}) calculated for TBA sequences

Sample	k_{diss} (s^{-1})
S26	3.4×10^{-3}
S38	3.8×10^{-3}

To confirm that the thermal dissociation of the anthracene photodimer restored the ability to fold into a quadruplex structure, sample **S26** was analysed again using CD spectroscopy. Figure 20 displays the CD spectra taken before and after thermal reversion of **S26-D$_1$** in the presence of K^+ and the results clearly show that the characteristic bands have increased in intensity following the thermal dissociation of the photodimer, in a similar fashion to Figure 14. This result confirms that the removal of the photodimer has removed the restriction placed on the oligonucleotide, allowing the quadruplex structure to reform.

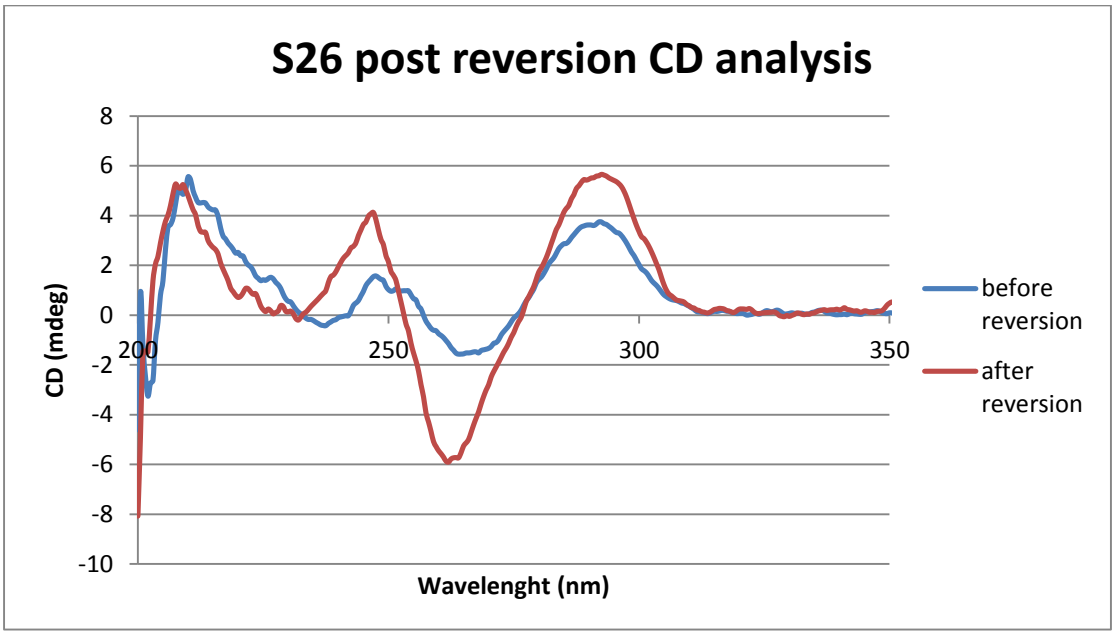


*Figure 20. CD spectra recorded before and after the thermal reversion of **S26-D$_1$**. An increase in the relative CD bands indicate the re-formation of a quadruplex structure (both spectra recorded in the presence of K^+).*

5.8 Potassium Sensing by Fluorescence Spectroscopy

A number of previous studies such as those of Takenaka and co-workers have utilised quadruplex forming oligonucleotides as potassium sensors with the use of fluorescent reporter groups.[6] Anthracene possesses fluorescent properties that are affected by changes within the local environment in which it exists. This property has already been exploited by the Tucker group to sense single nucleotide polymorphisms (SNPs) within short oligonucleotide sequences.[38] As mentioned in Section 5.5.1 the presence of potassium was found to affect the photodimerisation rate of group 4 sequences. Next it was decided to study potassium sensing by monitoring changes to the fluorescence spectrum of the quadruplex sequences.

The graphs in Figure 22 display the fluorescence results from sequences **S26**, **S30** and **S38** respectively, each of which utilise the same anthracene linker type (6D). From these results it is clear that some sequences display greater changes to their spectra than others upon the addition of potassium, with sequence **S38** providing the greatest change with a 40% decrease in emission at 430 nm (Figure 22C). In contrast sequence **S30** displayed almost no change in emission upon potassium addition, whereas there was a small decrease for sequence **S26** of 11%, despite confirmed quadruplex folding analysed through CD spectroscopy. These changes are most likely attributed to changes in static quenching, in particular with sequence **S38** where the anthracene is placed on opposite termini of the oligonucleotide, leaving it relatively exposed.[39] However upon folding both anthracene tags would come into close proximity with the lower G-tetrad of the quadruplex structure allowing it to π-stack with the lower G-tetrad (Figure 21). In

this environment, an increase in static quenching with neighbouring bases would be expected, leading to the observed decrease in the fluorescent emission intensity. A similar observation was made by Kim and co-workers who, in a similar strategy, modified the TBA quadruplex sequence with a single pyrene fluorophore to detect changes in the nucleic acid structure.[40] They too found that an oligonucleotide labelled with pyrene on the 5' end resulted in changes in the fluorescence spectra on the addition of potassium due to 'end-stacking'. The anthracene tags within **S26** and **S38** are located within the main body of the sequence and therefore would not be expected to undergo such a dramatic change in environment. The increased sensitivity of **S38** towards potassium binding makes it the most likely candidate for a potassium sensor out of three samples analysed.

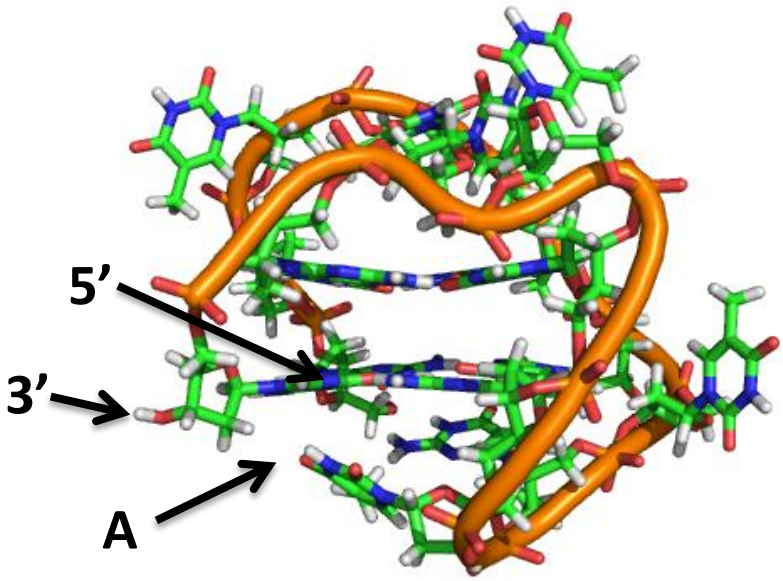


*Figure 21. A model of the TBA quadruplex produced from the 1RDE pdb file illustrating the position (**A**) most likely adopted by the anthracene tags in sequence **S38**.*

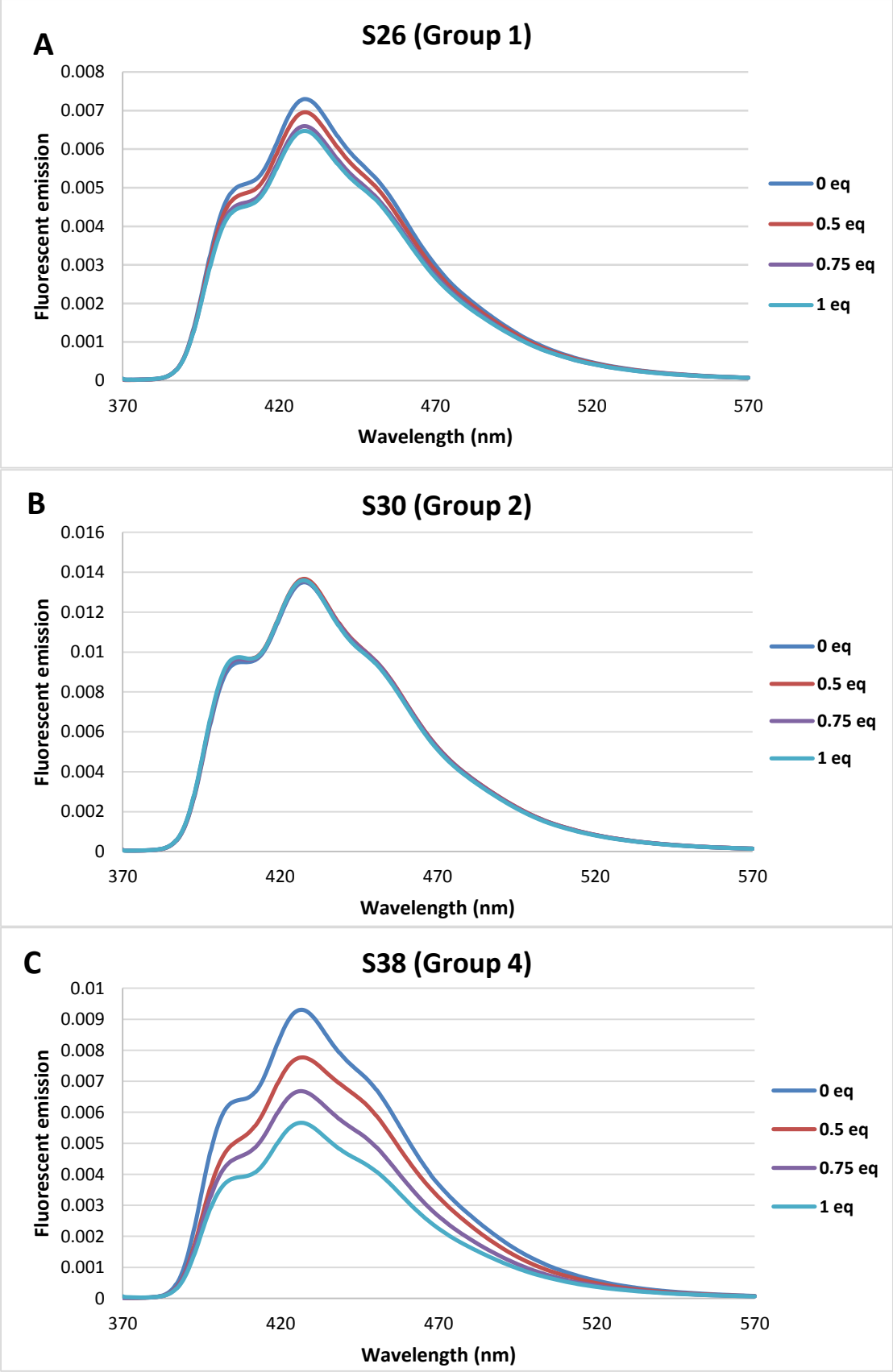


*Figure 22. (A-C) Changes to the fluorescence spectra for sequences **S26**, **S30** and **S38** respectively upon addition of potassium. (1mM, ambient temp., λ_{ex} – 360 nm).*

5.9 Thrombin Binding

The interaction between the TBA quadruplex structure and the thrombin protein and its role as an inhibitor has been investigated by a number of groups in recent years. The focus of the Sica group has been to model DNA/protein interactions through high resolution crystallographic structural characterisation experiments in order to gain an understanding of the key interactions that are involved.[41] In doing so they developed a crystal structure that has been made available through the protein data bank (PDB), which shows the protruding TT loops of the TBA sequence acting as a pincer that locks onto the external exosite of the protein (Figure 23A). The binding is achieved through a combination of hydrophobic and H-bonding interactions, which rely crucially on the folded structure of TBA being present. Therefore sequences within Group 1 where folding was successfully inhibited upon photodimerisation were expected to inhibit thrombin binding to a certain extent. However one consideration to note is that Group 1 sequences have a modified external loop (loop 3), where the anthracene group has been incorporated and therefore are missing one thymine nucleobase compared to the unmodified sequence. It was thought that this modification could have an impact on how the modified TBA would interact with the thrombin protein prior to photodimerisation, producing less predictable binding results.

Other research by Giancola, which focused on investigating the energetics of TBA/Thrombin binding through isothermal titration calorimetry (ITC), indicated that the binding stoichiometry of this interaction was 1:2.[42] Their model of the interaction depicts the TBA quadruplex sandwiched between two thrombin proteins (Figure 23B), where opposite faces of the quadruplex interact with different binding sites on the protein – exosite I (green) and exosite II (cyan).

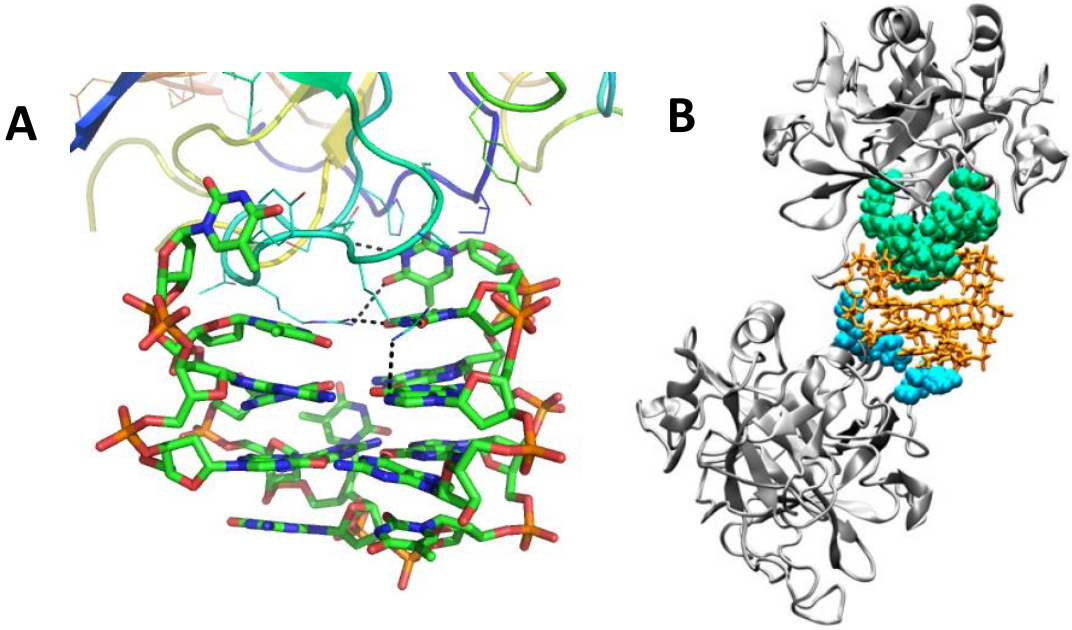


Figure 23. (A) Model illustrating some of the interactions that make TBA/thrombin binding possible (black dashed lines). (B) Earlier research demonstrated that TBA binds to thrombin in a 1 : 2 ratio to exosite I (green) and exosite II (cyan). Image taken from ref. 39.

The secondary thrombin protein binds to the thrombin through similar hydrophobic and polar interactions, except in this instance the interaction involves the opposite G-tetrad and the central TGT loop of the TBA sequence. Therefore additional sequences from Groups 2 and 4 could also regulate the binding of thrombin through photodimerisation, as the anthracene tags within these sequences are placed in locations that could play a key role in disrupting the second thrombin/TBA interaction.

A series of gel electrophoresis experiments were developed in order to investigate whether the dimerisation of the modified TBA sequence had an effect on the binding affinity towards thrombin (Sigma Aldrich – T6884). The tests were carried out under native conditions using a potassium salt within the gel and buffer in order to promote quadruplex formation.[43] A 20% polyacrylamide gel was used, which allowed the TBA sequence to migrate but still retain the thrombin protein within the sample well due to their different orders of magnitude in mass.[44] To assess the binding affinity of each sequence, a series of samples were made where the ratio of thrombin : **S26** was gradually increased from 5 : 1 to 1 : 4 in set increments (Table 5). The samples were mixed and incubated to aid binding before being placed in adjacent wells on the gel.

During the experiment the TBA quadruplex was found to bind to the thrombin, which due to the nature of the size of the protein, was retained within the sample well. Once the ratio of **S26** was raised to a point where there becomes an excess of unbound DNA, the unbound material then migrates onto the gel plate, allowing it to be visualised by imaging techniques (SYBR Gold staining). Figure 24A displays the gel image of the starting material **S26** paired with thrombin, whereas Figure 24B displays the gel image produced from pairing the photoproduct **S26-D$_1$** with thrombin. A comparison of these images would suggest that the photodimerisation of the anthracene tags within the **S26** sequence has indeed affected the binding towards the thrombin protein to some degree, as the DNA bands for the photoproduct within Figure 24B appear in more lanes. When the binding becomes weaker due to photodimerisation there is a greater excess of the unbound TBA oligonucleotide at lower sample ratios, which migrates onto the gel

and shows up through imaging. Conversely the starting material, which is more tightly bound by the protein, is retained within sample well until the amount of DNA is increased further, accounting for the difference in the two images.

Table 5. List of sample wells and corresponding sample ratios used for gel electrophoresis analysis.

Sample Well	Ratio (Thrombin : **S26**)
1	5 : 1
2	4 : 1
3	3 : 1
4	2 : 1
5	1 : 1
6	1 : 2
7	1 : 4

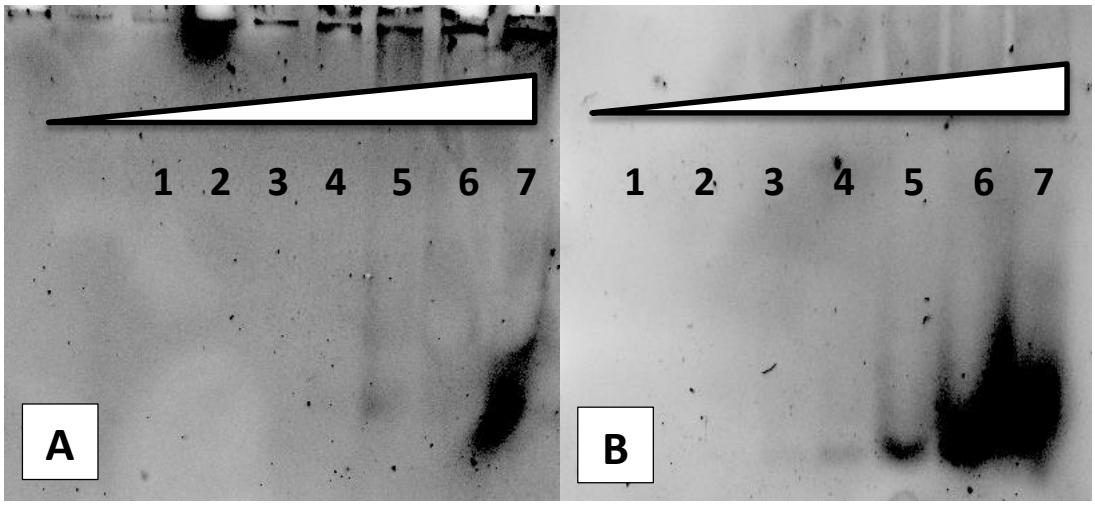


*Figure 24. Thrombin binding studies through gel electrophoresis using samples **S26** (A) and **S26-D$_1$** (B). (20% native-PAGE, 100mM KCl, 1 x TB, 19.5 hour run time, 100V) Images produced through SYBR gold staining and transillumination.*

Although this technique has provided some interesting observations, the results are only preliminary and will require further testing, including some control studies, in order to confirm that the interpretation of the results is correct. It is hoped that in the future the method can also be performed on a gradient gel, which

would allow visualisation of the bound TBA quadruplex on the gel, providing an even better indication of the change in binding towards thrombin.[45]

While the results from gel electrophoresis studies were informative they are however not quantitative. Therefore attempts were made to utilise isothermal titration calorimetry (ITC) in order to obtain more detailed thermodynamic data, which would give further insight into the effect anthracene photodimerisation has on TBA/thrombin binding.[46,47] Unfortunately the technique could not be perfected in the time available and the data is therefore not included within this discussion. However, future attempts to perfect this technique are planned as the ITC technique is considered an optimal method to get a quantitative indication of the changes within this DNA/protein interaction.

5.10 Conclusions and Further Work

The thrombin binding aptamer (TBA) quadruplex has proven to be an interesting, albeit challenging target in which to modify with anthracene tags. T_m and CD analysis has proven that the formation of the quadruplex secondary structure can be manipulated through anthracene photodimerisation, providing the correct tagging strategy is utilised. The best results obtained were for sequences within Group 1, where anthracene groups are placed in different loop segments within the quadruplex sequence. The formation of a quadruplex structure is very dependent on the loop portions of the sequence folding correctly to allow the guanine nucleobases to align into a tetramer, and this particular method of modification seems to have been the most effective in disrupting that folding. Further work to fully investigate the interaction of modified TBA sequences with

thrombin is still to be completed, including isothermal titration calorimetry and improved methods of gel electrophoresis.

Although use of the TBA sequence for this investigation has had certain benefits since it has a single polymorph and has a known interaction with thrombin, the method of incorporating anthracene could in theory be applied to larger quadruplex sequences containing more than two G-tetrads. The composition of unimolecular quadruplex sequences are essentially the same and can be generally described as $G_nX_aG_nX_bG_nX_cG_n$ (n = number of G-tetrads, X_a, X_b & X_c can be any combination of residues that will form the loops). Therefore anthracene photodimerisation could potentially be used to modulate a range of important processes that are associated with the formation of a unimolecular quadruplex.

5.11 References

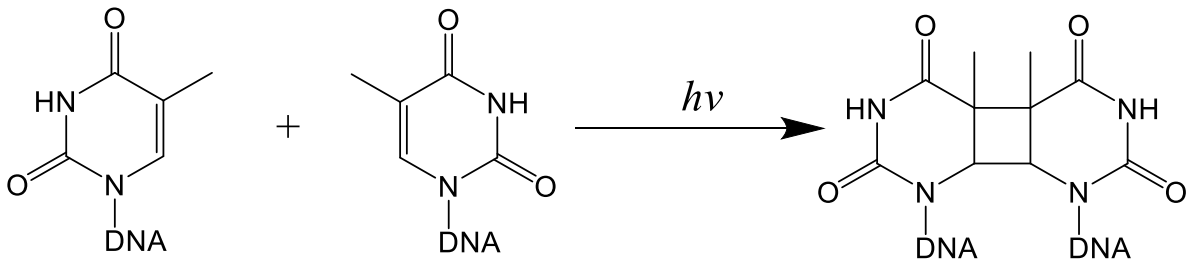
1. Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S., *Nucleic Acids Res.*, **2006**, *34*, 5402-5415.
2. Todd, A. K.; Johnston, M.; Neidle, S., *Nucleic Acids Res.*, **2005**, *33*, 2901-2907.
3. Gellert,M.; Lipsett, M. N.; Davies, D. R., *Proc. Natl. Acad. Sci. USA*, **1962**, *48*, 2013-2018.
4. Davis, J. T., *Angew. Chem. Int. Ed.*, **2004**, *43*, 668-698.
5. Guéron, M.; Leroy, J.-L., *Curr. Op. Struct. Biol.*, **2000**, *10*, 326-331
6. Nagatoishi, S.; Nojima, T.; Juskowiak, B.; Takenaka, S., *Angew. Chem. Int. Ed.*, **2005**, *44*, 5067-5070.
7. Li, J. J.; Tan. W., *Nano. Lett.*, **2002**, *2*, 315-318.
8. Sidorov, V.; Kotch, F. W.; El-Kouedi, M.; Davis, J. T., *Chem. Commun.*, **2000**, 2369-2370.
9. Blackburn, E. H., *Nature*, **1991**, *350*, 569-573.
10. Zahler, A. M.; Williamson, J. R.; Cech,T. R.; Prescott, D. M., *Nature*, **1991**, *350* 718-720.
11. Neidle, S., *FEBS Journal*, **2010**, *277*, 1118-1125.
12. Monchaud, D.; Teulade-Fichou, M.-P., *Org. Biomol. Chem.*, **2008**, *6*, 627-636.
13. Mergny, J.-L.; Lacroix, L.; Teulade-Fichou, M.-P.; Hounsou, C.; Guittat, L.; Hoarau, M.; Arimondo, P. B.; Vigneron, J.-P.; Lehn, J.-M.; Riou, J.-F.; Garestier, T.; Hélène, C., *Proc. Natl. Acad. Sci. USA*, **2001**, *98*, 3062-3067.
14. Sun, D.; Thompson, B.; Cathers, B. E.; Salazar, M.; Kerwin, S. M.; Trent, J. O.; Jenkins, T. C.; Neidle, S.; Hurley, L. H., *J. Med. Chem.*, **1997**, *40*, 2113-2116.
15. Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H., *Proc. Natl. Acad. Sci. USA*, **2002**, *99*, 11593-11598.
16. Han, F. X.; Wheelhouse, R. T.; Hurley, L. H., *J. Am. Chem. Soc.*, **1999**, *121*, 3561-3570.
17. Heckel, A.; Buff, M. C. R.; Raddatz, M.-S. L.; Müller, J.; Pötzsch, B.; Mayer, G., *Angew. Chem. Int. Ed.*, **2006**, *45*, 6748-6750.
18. Mayer, G.; Kröck, L.; Mikat, V.; Engeser, M.; Heckel, A., *ChemBioChem*, **2005**, *6*, 1966-1970.
19. Ogasawara, S.; Maeda, M., *Angew. Chem. Int. Ed.*, **2009**, *48*, 6671-6674.
20. Wang, X.; Huang, J.; Zhou, Y.; Yan, S.; Weng, X.; Wu, X.; Deng, M.; Zhou, X., *Angew. Chem. Int. Ed.*, **2010**, *49*, 5305-5309.
21. Gaynutdinov, T. I.; Neumann, R. D.; Panyutin, I. G., *Nucleic Acids Res.*, **2008**, *36*, 4079-4087.
22. Dailey, M. M.; Miller, M. C.; Bates, P. J.; Lane, A. N.; Trent, J. O., *Nucleic Acids Res.*, **2010**, *38*, 4877-4888.
23. Hayashida, H.; Paczesny, J.; Juskowiak, B.; Takenaka, S., *Bioorg. Med. Chem.*, **2008**, *16*, 9871-9881.

24. Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J., *Nature*, **1992**, *355*, 564-566.
25. Macaya, R. F.; Schultze, P.; Smith, F. W.; Roe, J. A.; Feigon, J., *Proc. Nat. Acad. Sci. USA*, **1993**, *90*, 3745-3749.
26. Schultze, P.; Macaya, R. F.; Feigon, J., *J. Mol. Biol.*, **1994**, *235*, 1532-1547.
27. Mao, X.; Marky, L. A.; Gmeiner, W. H., *J. Biomol. Struct. Dyn.*, **2004**, *22*, 25-33.
28. Padmanabhan, K.; Padmanabhan, K. P.; Ferrara, J. D.; Sadler, J. E.; Tulinsky, A., *J. Bio. Chem.*, **1993**, *268*, 17651-17654.
29. Davie, E. W.; Fujikawa, K.; Kisiel, W., *Biochemistry*, **1991**, *30*, 10363-10370.
30. Kraus, I. R.; Merlino, A.; Giancola, C.; Randazzo, A.; Mazzarella, L.; Sica, F., *Nucleic Acids Res.*, **2011**, *39*, 7858-7867.
31. Pasternak, A.; Hernandez, F. J.; Rasmussen, L. M.; Vester, B.; Wengel, J., *Nucleic Acids Res.*, **2011**, *39*, 1155-1164.
32. Lane, A. N.; Chaires, J. B.; Gray, R. D.; Trent, J. O.; *Nucleic Acids Res.*, **2008**, *36*, 5482-5515.
33. Olsen, C. M.; Lee, H.-T.; Marky, L. A., *J. Phys. Chem. B*, **2009**, *113*, 2587-2595.
34. Zhang, A. Q., Balasubramanian, S., *J. Am. Chem. Soc.*, **2012**, *134*, 19297-19308.
35. Gray, D. M.; Wen, J.-D.; Gray, C. W.; Repges, R.; Repges, C.; Raab, G.; Fleischhauer, J., *Chirality*, **2008**, *20*, 431-440.
36. Chaires, J. B.; Graves, D., *Quadruplex Nucleic Acids*, Springer Publishing, Berlin, **2013**, 67-86.
37. Nagatoishi, S.; Tanaka, Y.; Tsumoto, K., *Biochem. And Biophys. Res. Comm.*, **2007**, *354*, 837-838.
38. Duprey, J.-L. H. A.; Zhao, Z.; Bassani, D. M.; Manchester, J.; Vyle, J. S.; Tucker, J. H. R., *Chem. Commun.*, **2011**, *47*, 6629-6631.
39. Lacowicz, J. R., *Principles of Flourescence Spectroscopy*, 3rd ed., Springer Publishing, New York, **2006**, Chapter 6, 205-235.
40. Seo, Y. J.; Lee, I. J.; Yi, J. W.; Kim, B. H., *Chem. Commun.*, **2007**, 2817-2819.
41. Kraus, I. R.; Merlino, A.; Randazzo, A.; Novellino, E.; Mazzarella, L.; Sica, F., *Nucleic Acids Res.*, **2012**, *40*, 8119-8128.
42. Pagano, B.; Martino, L.; Randazzo, A.; Giancola, C., *Biophys. J.*, **2008**, *94*, 562-569.
43. Sosic, A.; Meneghello, A.; Cretaio, E.; Gatto, B., *Sensors*, **2011**, *11*, 9426-9441.
44. Garner, M. M.; Revzin, *Nucleic Acids Res.*, **1981**, *9*, 3047-3060.
45. Hames, B. D., *Gel Electrophoresis of Proteins, a Practical Approach*, **1998**, Oxford University Press.
46. Ladbury, J. E.; Chowdhry, B. Z., *Chem. & Biol.*, **1996**, *3*, 791-801.
47. Lin, P.-H.; Chen, R.-H.; Lee, C.-H.; Chang, Y.; Chen, C.-S.; Chen, W.-Y., *Colloids and Surfaces B: Biointerfaces*, **2011**, *88*, 552-558.

Chapter 6: Inter-Strand Cross-Linking

6.1 Introduction.

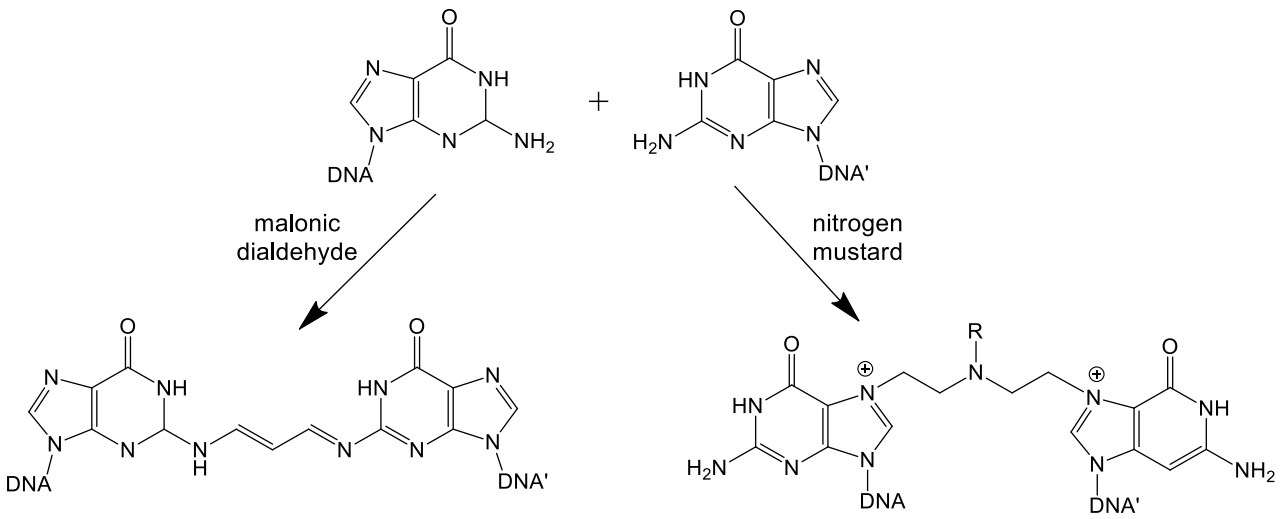
Within the past few years a number of human disorders have been discovered to be a result of direct damage to DNA structure, including a variety of cancers such as melanoma, carcinoma and lymphoma.[1] Defects within the repair pathways that are known to overcome these mutations within DNA can lead to an accumulation of damage, at which point a cell will cease to function normally and in some cases become cancerous.[2-3] Many of the preventable forms of DNA damage are caused by exogenous factors, meaning they originate from outside the body. These include UV light, X-Ray radiation and mutagenic compounds.[4-5] A common side effect of the absorption of UV light by DNA is the intra-strand cross-linking of adjacent cytosine or thymine nucleobases through covalent bonds (Scheme 1).[6] This type of damage can inhibit DNA replication and is the primary cause of melanomas, which are forms of skin cancer.[7] UV light induced-damage to DNA is also the main reason why optical filters are essential for the application of photochemical reactions to DNA based systems, targeting only the desired reactive species such as anthracene or azobenzene.



Scheme 1. Cross-linking of adjacent thymine nucleobases due to the absorption of UV-B light radiation.

6.1.1 Inter-Strand Cross-Link

Another important type of DNA damage is what is known as inter-strand cross-linking (ICL), where a reactive species creates covalent linkages to nucleobases on opposite strands of a DNA helix.[8] An inter-strand cross-link prevents a double stranded helix from dissociating into its single stranded components, therefore blocking vital transcription and metabolism processes and is therefore one of the most cytotoxic lesions currently known.[9] They arise from the reaction with bifunctional electrophilic compounds that are either developed in the body, such as malonic dialdehyde (a product of lipid peroxidation), or exogenous compounds such as chlorambucil, an anti-cancer drug, (Scheme 2).[10] The main focus of the research into ICL forming compounds is therefore for their use as anti-tumour agents, due to their ability to induce apoptosis of malignant tumour cells through the resulting cross-link.[11]



Scheme 2. Cross-linking of guanine nucleobases through an endogenous compound (malonic dialdehyde) and an exogenous compound (A nitrogen mustard such as chlorambucil).

Another leading anti-tumour drug, cisplatin, is also known to generate cross-linked adducts, 5-10% of which are inter-stranded and is widely used to treat a variety of cancers.[12] However, the natural response to the presence of cross-links within a sequence is to remove them through enzymatic repair pathways to regenerate the natural DNA, which therefore leads to a reduction in the effectiveness of these anti-tumour drugs.[13] The main issue involved with the research of ICL type anti-tumour drugs is that the mechanism by which they are repaired within the cell is as yet unknown. In addition to this there are further complications in that the study of ICL repair is made difficult due to the lack of model substrates on which to perform the mechanistic studies. This is mainly due to difficulty in isolating a singular cross-linked product in sufficient yields from biological material.

5.1.2 Synthetic Inter-Strand Cross-Linking

Research is currently underway in an effort to achieve artificial initiation of inter-strand cross-links using reactive species built into a synthetic oligonucleotide. One such project currently being led by Madder and co-workers involves the synthesis of a modified nucleotide with an incorporated furan moiety, which once incorporated into an oligonucleotide is capable of cross-linking to the complimentary strand.[14-15] The cross-link is initiated through oxidation of the furan moiety with *N*-bromosuccinimide to give a reactive dicarbonyl that immediately reacts with the complementary base, (Figure 1).[16] The development of artificial DNA cross-links through modified oligonucleotides has proved a vast improvement over other techniques involving bifunctional alkylating agents, which typically generate low yields of the cross-linked product (10%). An additional advantage is the selectivity of the technique, which generates a single cross-

linked product that can easily be characterised, an important trait that is needed for the effective study of repair enzymes. The ultimate goal of this research is to develop methods to better understand ICL repair pathways, using the knowledge gained to improve the clinical effectiveness and to investigate the possible side effects of existing anti-tumour drugs.

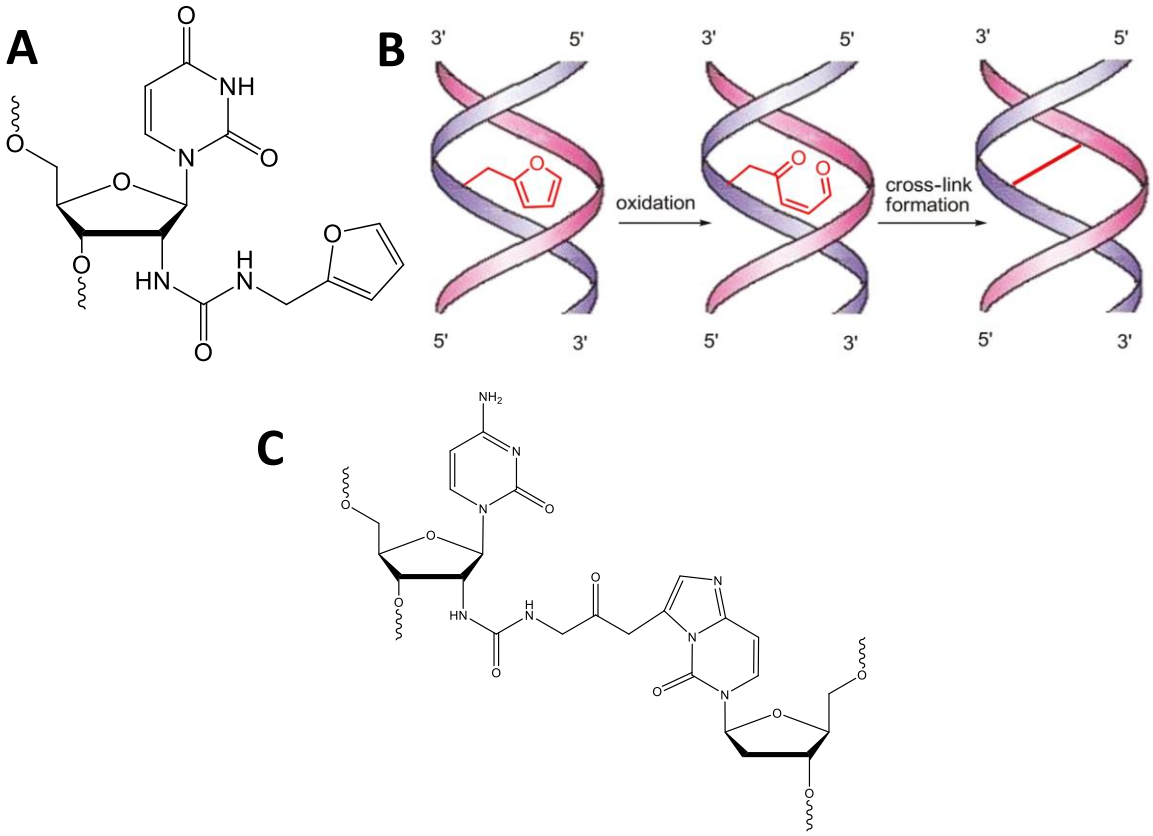
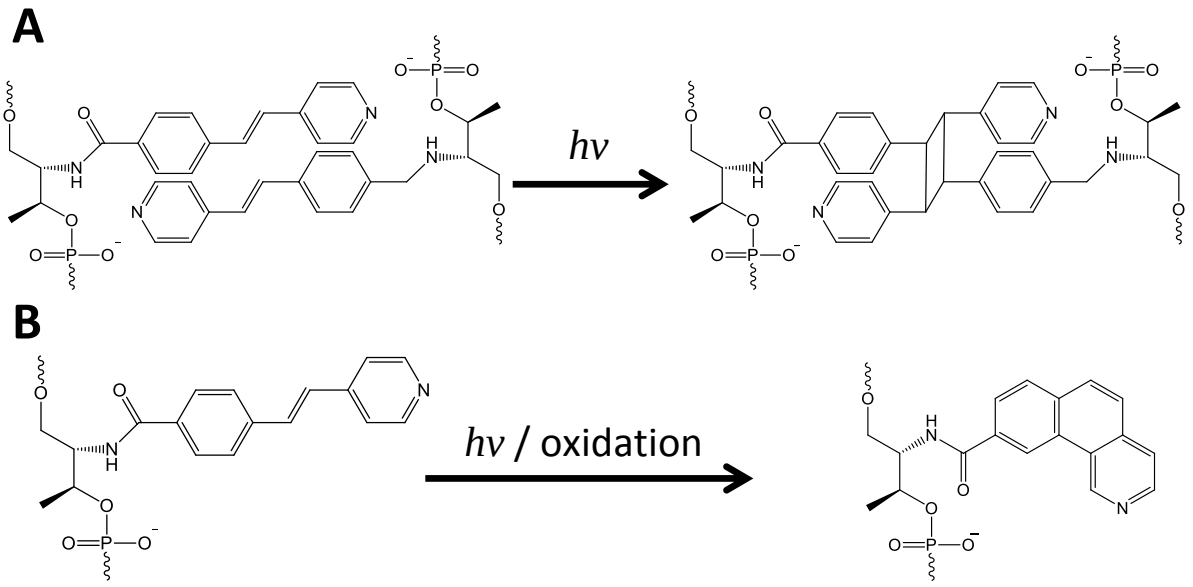


Figure 1. (A) Modified nucleotide with furan moiety. (B) Initial formation of dicarbonyl with N-bromosuccinimide followed by cross-link formation with the complementary strand. Image taken from ref. 14 (C) Structure of the resulting cross-link to a cytosine on the adjacent oligonucleotide.

By placing anthracene tags onto opposite positions of two complementary oligonucleotides it may be possible to create the same effect of cross-linking the modified strands through photodimerisation. Since anthracene is a photochromic compound the process would also be reversible and could be used to study repair pathways in a similar fashion to the research of Madder *et al*. Additionaly the anthracene dimerisation is driven

by light, therefore the cross-linking could be performed *in vivo* allowing much more flexibility over previous techniques. Researchers would be able to control where and when a cross-link is created by targeting certain areas within a cell or tissue using UV light to initiate the anthracene cross-link, this could be useful for when the research focuses on targeting tumorous cells over healthy cells.



Scheme 3. (A) [2+2] photo-cycloaddition of a stilbazole derivative placed on opposite to create an interstrand cross-link (ICL). (B) Light driven trans/cis isomerisation and subsequent oxidation to benzoquinoline, only in the single stranded or monomeric state.

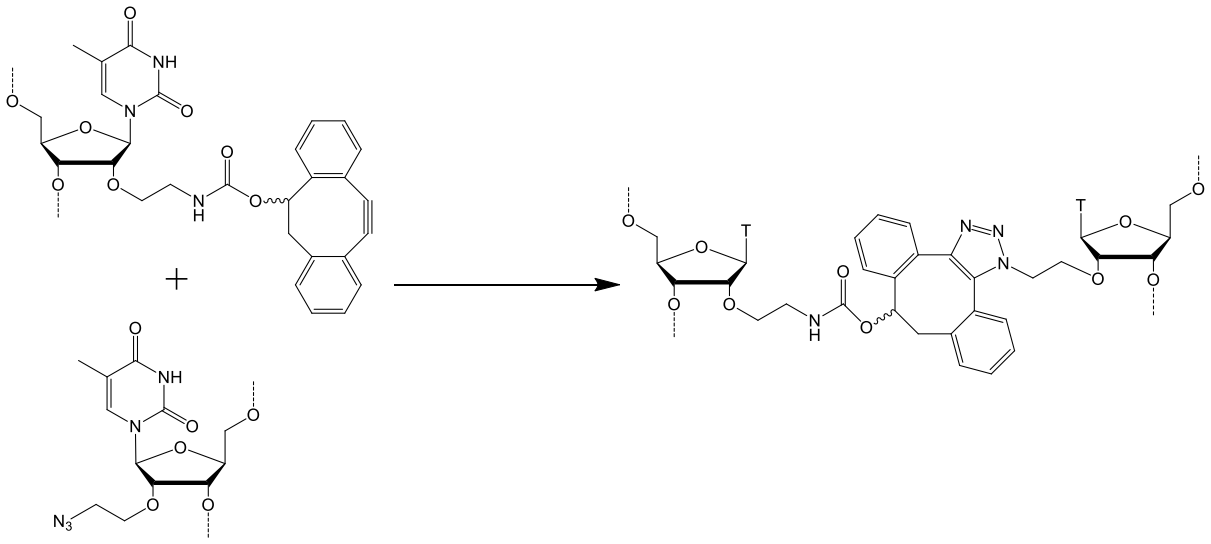
In very recent research, the use of photochromic compounds has already been used to create a reversible cross-link between modified complementary oligonucleotides. Asanuma and co-workers developed a stilbazole derivative that when placed within adjacent sites on the complementary strands, can undergo a photo-induced [2+2] cycloaddition reaction, (Scheme 3A).[17-18] When irradiated in either its monmeric form, or when incorporated into a single stranded oligonucleotide, the stilbazole moeity normally

undergoes a number of undesired side reactions, including isomerisation to a benzoquinoline derivative. This involves the structure first isomerising to the *cis* configuration before being oxidised to form the three membered ring system (Scheme 3B). However within the environment of the duplex structure, the interactions with neighbouring bases restricts rotation about the double bond, leading to the formation of a single photoproduct, the cycloadduct.

The main focus of Asanuma's research was to develop techniques to improve the heat tolerance of a DNA-duplex for nanotechnology purposes, as was seen in the T_m data collected for the stilbazole photoproduct, which increased from 46.4 °C to >80 °C as a result of the inter-strand cross-link. Therefore cross-linking of duplex structures can help in the design of DNA based architectures and devices by improving their resistance to degradation in various environments. For example it has already been demonstrated that photoligation of self assembled oligonucleotides within a tiled array complex greatly aids in the structures stability, therefore aiding their nano-material based applications.[19] The technique also has the potential to be applied to three dimensional DNA architectures, rigidifying the framework of the nano structures once they are assembled and therefore enhacing their encapsulation properties.

Another form of artificial inter-strand cross-linking has recently been developed by Brown and co-workers through a form of copper free click chemistry known as ring strain-promoted azide-alkyne [3+2] cycloaddition (SPAAC). This method involved the synthesis of a dibenzocyclooctyne(DIBO)-tymidine monomer, which was incorporated into a series of oligonucleotides and paired with complementary strands containing an

azide label to initiate the cross-linking cycloaddition reaction (Scheme 4).[20] As this form of click chemistry does not require metal-ion catalysis the method has scope for *in vivo* applications, since Cu^{I} used in traditional copper catalysed cycloaddition reactions is known to be cytotoxic and can initiate DNA damage in the presence of oxygen. The resulting cross-linked duplexes provided an 18.5 – 28.5 °C inrease in T_m over the uncross-linked controls, with the best results achieved when the DIBO and azide modifications were attached through the 2'-oxygen of the ribose sugar, allowing them to cross-link across the minor groove. This technique has previously been utilised in templated DNA strand ligation as well as post synthetic labelling,[21] and in the light of this new research it has the potential for use in DNA repair studies and a range of other applications.



Scheme 4. Artificial cross-linking of oligonucleotides through ring strain-promoted azide-alkyne [3+2] cycloaddition (SPAAC).

6.2 Design and Synthesis of Cross-Linking Oligonucleotides

*Table 1. Sequence composition of strands synthesised for ICL studies including a control sequence (**T4**), which includes an abasic modification (**A_b**) in place of anthracene.*

Sequence	Linker	Sequence composition (5'-3')
S41	7D	TGGACTC**X**CTCAATG
S42	4D	
T2	7D	CATTGAG**X**GAGTCCA
T3	4D	
T4	-	CATTGAG**A_b**GAGTCCA

In order to develop an anthracene cross-linking technique that utilises anthracene photodimerisation, two types of oligonucleotide were synthesised, both complementary to each other and containing a single anthracene tag. The tags were designed to be placed within the sequence so that upon hybridization they would be located directly opposite each other, maximising their proximity and therefore the probability of successfully achieving photodimerisation, (Table 1). The sequence composition within these modified strands is exactly the same as those used in the study within Chapter 3 so that data could be more easily be correlated and compared. Two types of linker were used for this study; the 7D in **S41/T2** and the 4D in **S42/T3**. The longer *n* = 7 linker was chosen as it was expected that the anthracene groups would require the maximum amount of flexibility in order to bridge the gap between complimentary strands, with the *n* = 4 linker included as a comparison.

5'-TGGACTC**X**CTCAATG-3'
3'-ACCTGAG**X**GAGTTAC-3'

Figure 2. Cross-linking duplex with a single anthracene modification on the central position on each of the complementary sequences. Hybridisation brings the two anthracene groups into close proximity allowing photodimerisation and therefore formation of a cross-link to take place.

6.3 Results and discussion

6.3.1 Irradiation and Mass Spectrometric Analysis

Prior to irradiation a fluorescent titration experiment was carried out in order to test the luminescent properties of the **S41/T2** duplex. Due to the fact that the anthracene tags on opposing strands would be brought into close proximity with each other through hybridisation, it was expected that an excimer could form. Figure 2 demonstrates that this was indeed the case, with an increase in the longer wavelength emission observed upon complementary binding, which indicated the presence of excimers. As a control, a fluorescent spectrum was also obtained from a duplex **S41/T4** with an abasic site on the complementary strand, (Figure 3), which as expected, did not display any longer wavelength emission. The detection of excimer fluorescence is a good indication that the anthracene tags can both intercalate into the space between oligonucleotide base pairs, which would provide optimal conditions for photodimerisation and the creation of an interstrand cross-link.

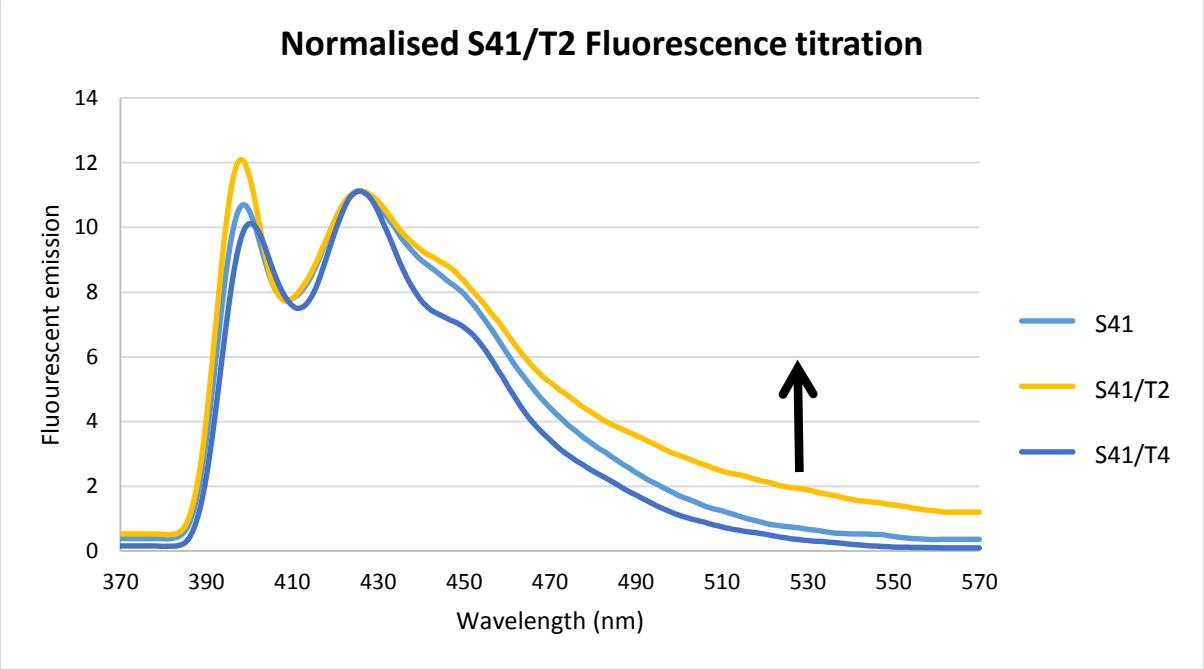


*Figure 3. Fluorescent titration of **S41** with increasing amounts of **T2**. Increase in the emission at longer wavelengths indicates the presence of excimers.*

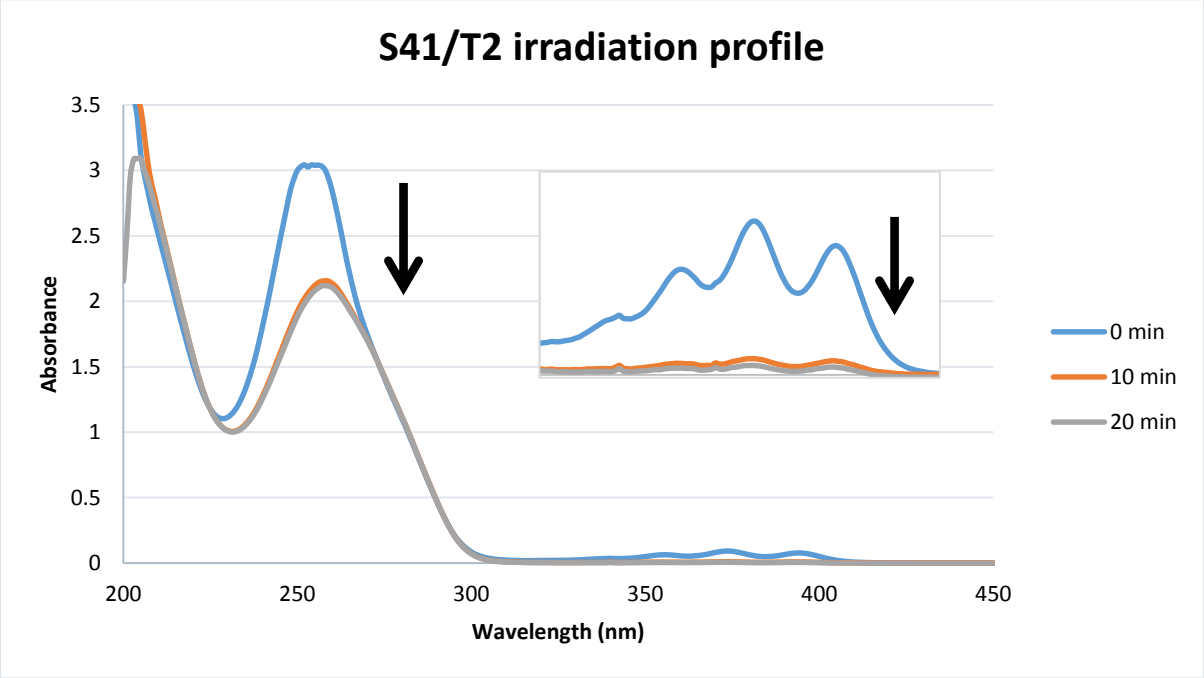


*Figure 4. UV analysis of ICL strands **S41/T2** through the process of irradiation (365 nm).*

To test the photochemical properties of the anthracene ICL strands, samples of **S41/T2** and **S42/T3** were mixed in equimolar ratios in Ar degassed, buffered solutions (20 μM

sample, 100 mM NaCl, 10 mM phosphate, pH 7) before being subjected to UV irradiation, which was filtered at 365 nm. Each sample was monitored through UV spectroscopy and both were found to be reasonably reactive, with the majority of both reactions occurring within 10 minutes and reaching completion within 20 minutes. Figure 4 displays a UV spectrum of the sample taken at set intervals during the reaction, where a significant decrease is observed in the anthracene absorption bands at 260 and 365 nm, which indicates conversion to the cross-linked photoproduct.

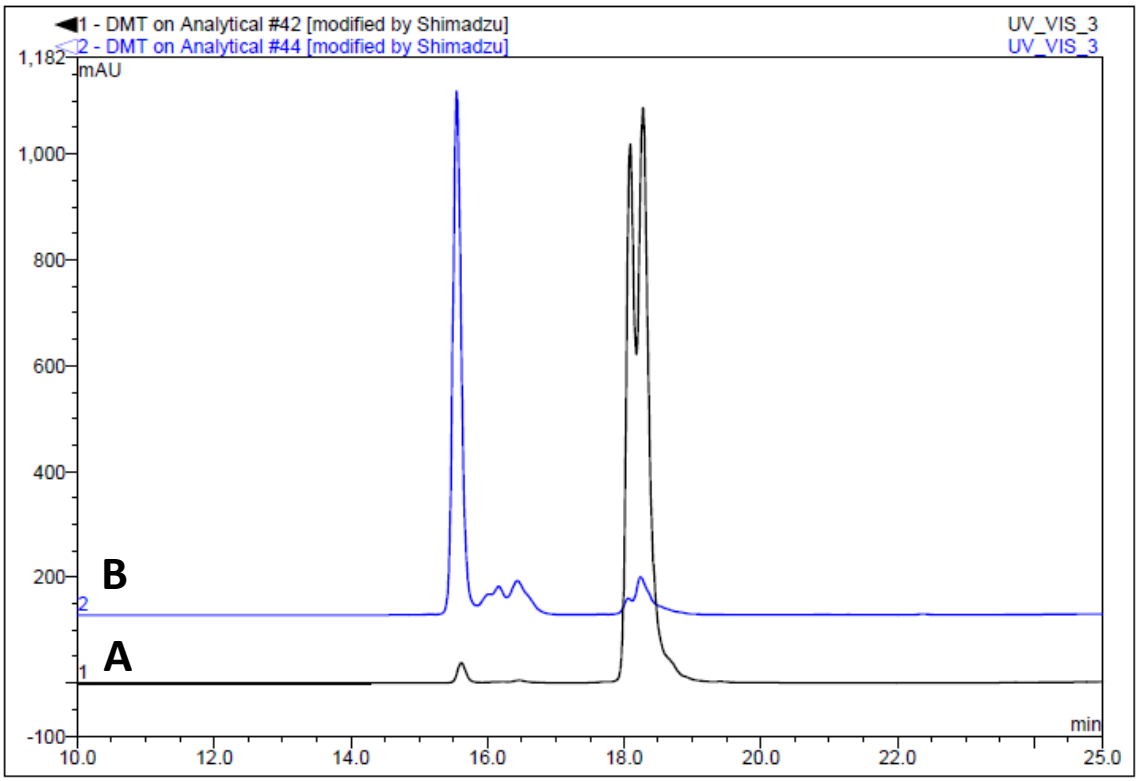
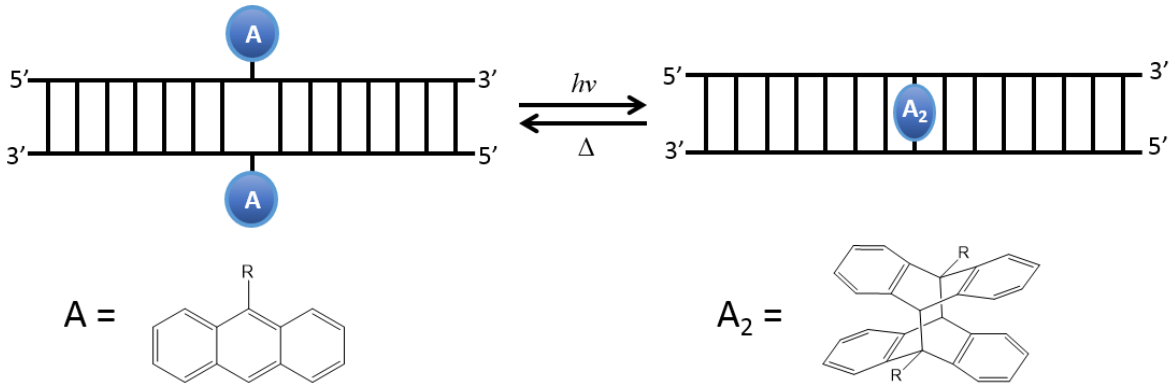


*Figure 5. (A) HPLC chromatogram of **S41/T2** prior to irradiation. (B) Resulting HPLC chromatogram of **S41/T2** following UV irradiation. Initial starting material peaks have disappeared, replaced by a singular cross-linked photoproduct peak.*

The samples were also monitored through reversed phase analytical HPLC to determine the composition of the reaction mixture before and after cross-linking of the complementary strands. Figure 5A displays the analytical HPLC chromatogram of the

S41/T2 sample prior to irradiation. Any duplex formation between the two complementary strands is removed under the conditions of HPLC analysis, which is why two peaks are observed. However due to their identical length and modification their retention times are very similar (18.09 and 18.27 min) and almost merge into a single peak. Upon conversion to the ICL photoproduct these two peaks disappeared and were replaced with one major product peak with a lower retention time (15.56 min). A minor product was also observed; however this product was low yielding and could not be identified as a cross-linked product through other techniques such as mass spectrometry (Figure 5B). It is possible that this peak is another isomer of the anthracene dimer, and if the same trends as those found for intramolecular photodimerisation are followed (Chapter 3.), it is most likely that the minor product is in fact the HH isomer. This would explain why it could not be isolated and characterised as an ICL product as it would be expected to be much less stable and not able to withstand subsequent purification steps. The observation of a major peak that gives rise to a stable photoproduct would suggest that the angle of approach between anthracene tags on complementary strands overwhelmingly favours the HT product, providing the cleanest photodimerisation reaction yet seen in this project.



Scheme 5. A schematic illustrating the creation of anthracene inter-strand cross-link through photodimerisation of anthracene tags placed on complementary sequences.

Following HPLC purification of the new photoproduct, mass spectrometric analysis was carried out in order to determine whether cross-linking had been successfully achieved. The mass observed in the resulting spectrum indeed matched what was predicted as the summation of masses of sequences **S41** and **T2**, confirming that the cross-linked duplex had formed (Figure 6).

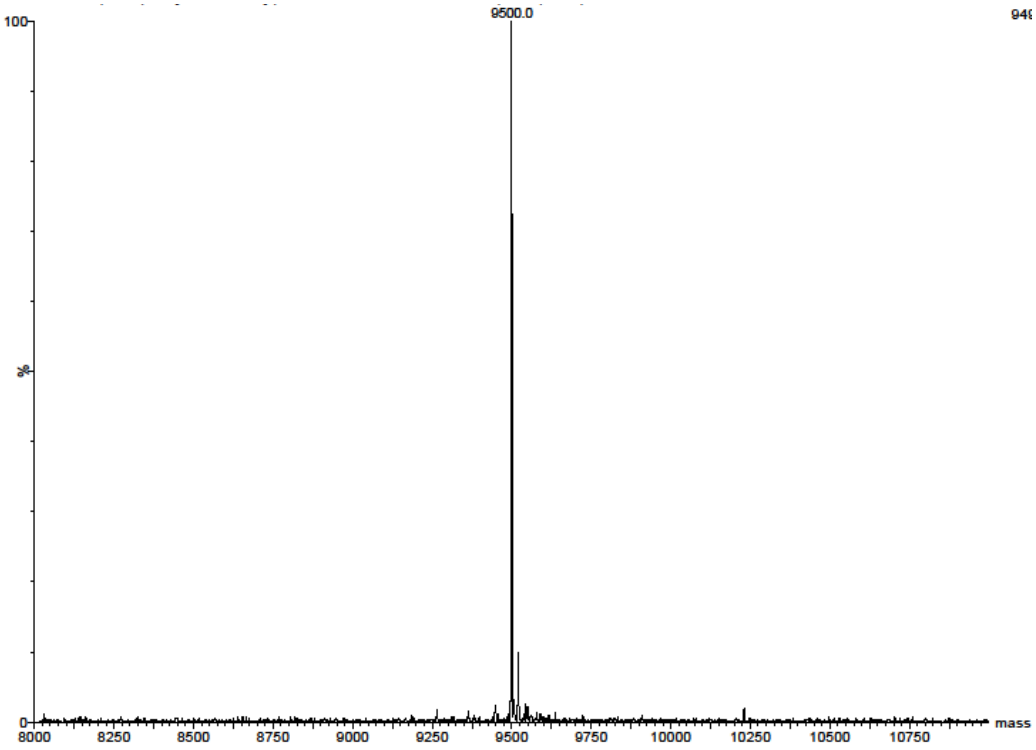


*Figure 6. ES- mass spectrum of **T2/S41** (n = 7) photoproduct peak.*

6.3.2 T_m analysis.

Table 2. T_m data collected for cross-linked sequences

Complex	Starting Material T_m	Cross-linked product T_m
S41/T2	53.5	> 70
S42/T3	52	> 70

T_m analysis of the cross-linked photoproducts was carried out to determine what effect interstrand photodimerisation had on the stability of the duplex structure. The initial T_m

for the **S41/T2** duplex prior to irradiation was measured to be 53.5 °C, not too dissimilar to that for the corresponding unmodified duplex **S0/T0** (55 °C), indicating that the modification had not significantly disrupted hybridisation. In other words the removal of a base pair is compensated by anthracene intercalation and the resulting π-stacking interactions, as indicated in earlier control studies (Chapter 3.9). Following dimerisation, the T_m value was significantly affected for both systems, displaying significant increases over the non-dimerised starting material (Table 2). In fact the change was so dramatic that the sigmoidal transition produced by the dissociation of the duplex structure was not complete at 85 °C (Figure 7). Therefore the T_m values for the anthracene cross-linked sequences could not be accurately determined but the data would suggest that the value was at least greater than 70 °C, an increase of 16.5 °C over the staring material. This result is comparable with research by Asanuma and the stilbazole ICL, which also a gave T_m greater than 80 °C.[17]

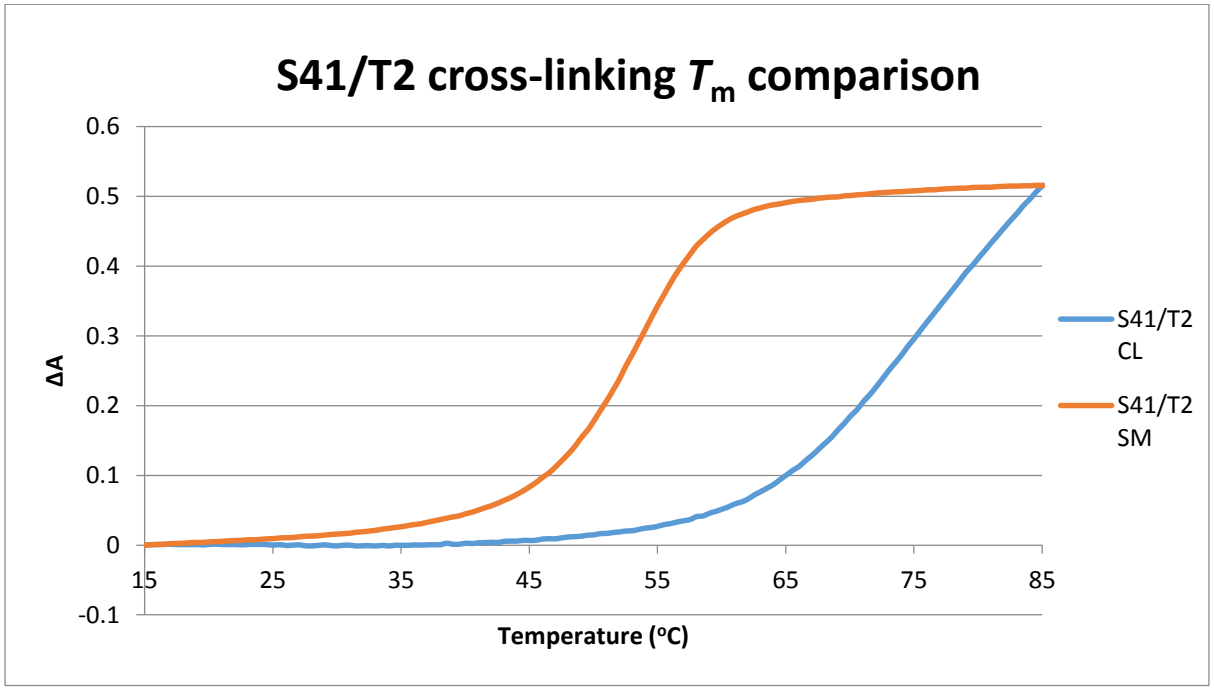


Figure 7. Comparison of the raw T_m data collected for the cross-linking sequences before and after irradiation. The anthracene cross-link has significantly stabilised the duplex, with the T_m transition going off the scale.

The high temperatures used to generate T_m data had a detrimental effect on the cross-linked product; though looking at the successive ramps from the raw T_m data (Figure 8), it was observed that the original T_m transition of the staring material duplex had reappeared. This would indicate that the anthracene cross-link had started to be removed through thermal reversion back to the starting material. In contrast to the intra-strand photodimerisation of anthracene (Chapter 3), the inter-strand cross-linked anthracene was much less stable to the thermal conditions used for the T_m analysis. This observation would suggest an increased structural strain placed on the anthracene dimer during strand dissociation at the elevated temperatures of the experiment, leading to the greater instability. The change in absorbance observed in the T_m experiment is therefore not solely due to the hyperchromicity effect of the duplex unwinding, but is also combined with the increased absorption that results from the reversion of the anthracene dimer. This may mean that any T_m data collected from these samples is slightly more inaccurate, nevertheless the change in absorbance occurs at a much higher temperature than that of the starting material and proves that the cross-link is in place, giving a duplex that is much more resistant to thermally induced base-pair dissociation than the non-cross-linked starting material. Many of the processes to which this technique could be applied to, such as transcription or branch migration, occur at lower temperatures, in a temperature range that the anthracene photodimer would be able to withstand.

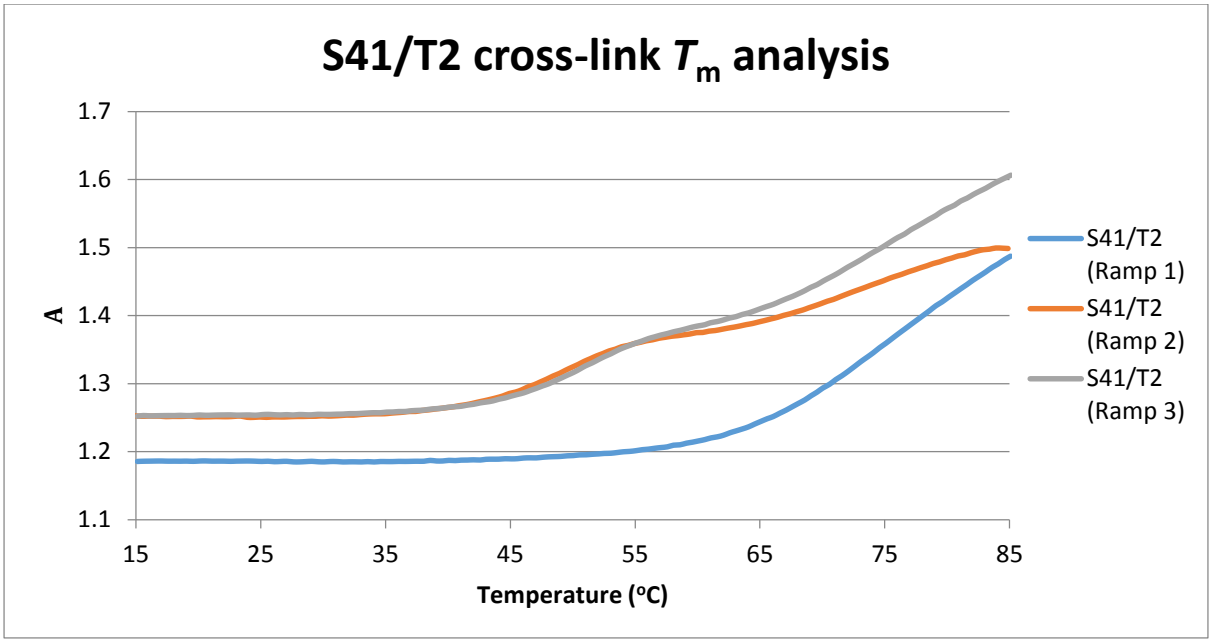


*Figure 8. Successive ramps collected for the cross-linked **S41/T2** duplex. After the first dissociative ramp (Ramp 1), the original T_m transition of the starting material reappears.*

An important factor that has a significant influence on the overall stability of a duplex is the presence of counterions, which neutralise the poly-anionic phosphate backbone of DNA. Lowering the salt concentration of a sample can lead to a decrease in the T_m as the oligonucleotides are no longer capable of binding effectively to form the stable double helix structure due to electrostatic repulsion.[22] However by pre-forming the duplex and initiating cross-linking, the resulting adduct may be resistant to changes in salt concentration. Figure 9 displays the T_m data for the cross-linked duplex **S41/T2-CL** and control duplex **S0/T0** using different concentrations of NaCl to monitor the affect it has on duplex stability. Whilst the T_m of the control duplex decreased considerably by 20 °C upon lowering the salt concentration, the effect on the cross-linked sample was much less, as shown by the small change the appearance of the T_m transition. Once again the transition measured under a lowered salt concentration did not display an upper limit, though the observed shift indicates that the ICL duplex is affected by these conditions to

some degree. The bases flanking the anthracene cross-link may not hydrogen bond as effectively under the lower salt concentration, however the cross-link seems to promote hydrogen bond formation despite these conditions, providing additional stability of ~35 °C over the control duplex.

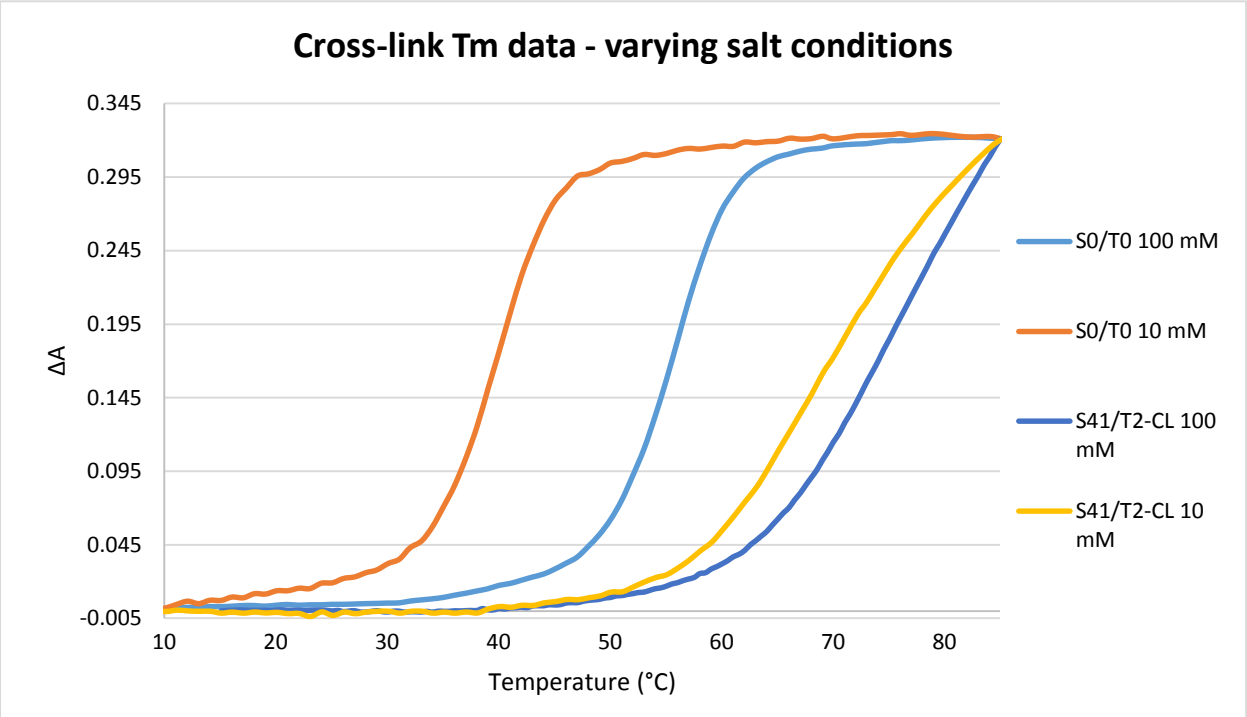


*Figure 9. T_m data of ICL duplex **S41/T2-CL** and control duplex **S0/T0** at varying salt concentrations (10 mM and 100 mM NaCl) to monitor effects on duplex stability. While the unmodified duplex **S0/T0** displays a dramatic drop in stability the cross-linked duplex **S41/T2-CL** exhibits a minor decrease.*

6.3.4 CD Analysis

As described in Chapter 3, CD analysis of the modified duplexes and their photoproducts is useful for identifying how the anthracene tags and their photodimers interact with the duplex structure. In the duplexes detailed here, both anthracene tags on the complementary sequences occupy the same site within neighbouring base pairs. Therefore it was of interest to observe what effect this would have on the induced

anthracene band within the CD spectrum. As shown in Figure 10 this is observed as a shoulder on the positive B-DNA band for the **S41**/**T2** duplex and is much more intense than that seen in earlier results (Section 3.6.2). The longer reach of the *n* = 7 linker would be expected to provide better flexibility to the anthracene groups and therefore allow

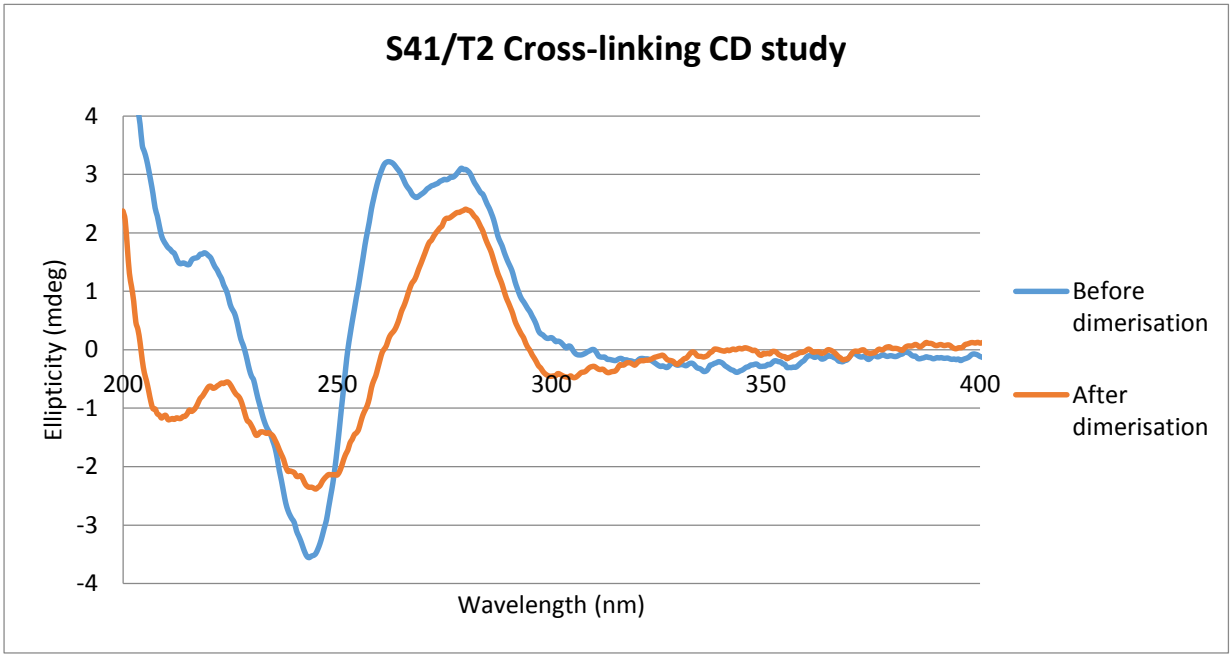


*Figure 10. CD analysis of n = 7 anthracene ICL samples (**S41** & **T2**) before and after dimerisation. The induced anthracene band disappears after dimerisation as well as small changes in the B-DNA bands indicating small changes to the duplex structure on formation of the cross-link.*

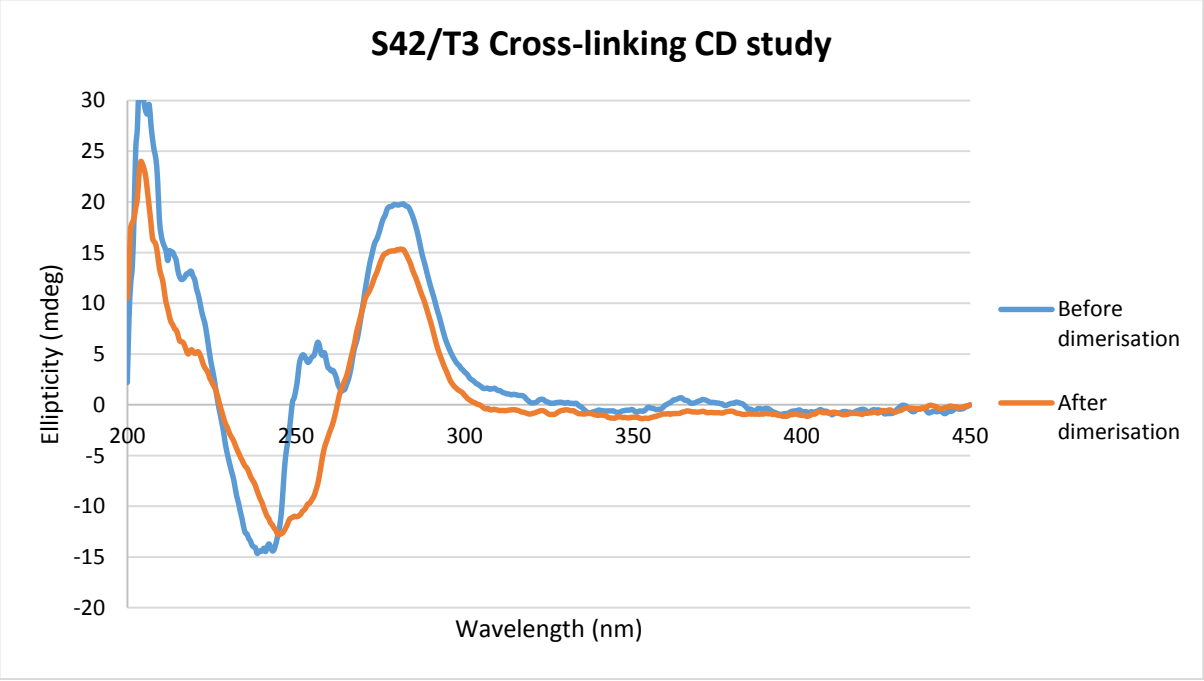


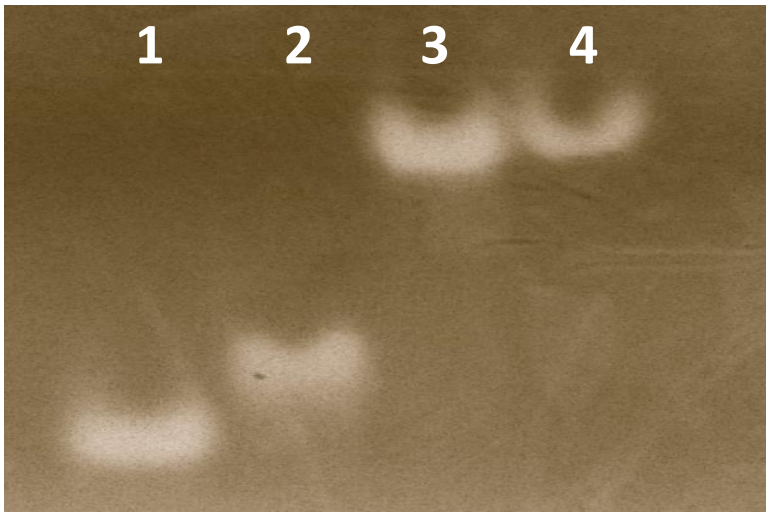
*Figure 11. CD analysis of n = 4 anthracene ICL samples (**S42** & **T3**).*

better overlap with neighbouring bases, which could explain why the band is more prominent than with the **S42/T3** duplex. With improved overlap the interactions between the anthracenes and the chiral base pair stack are enhanced, leading to the larger induced anthracene band. This result gives further support to the idea that both anthracene groups have successfully intercalated, which is consistent with the previous fluorescence studies, suggesting that the duplexes are ideally set up for the cross-linking photodimerisation reaction.

As expected, upon photodimerisation to form the cross-linked photoproduct, the induced anthracene band disappeared, in accordance with the anthracenes converting into the non-planar photodimer. However the important observation for both **S41/T2** and **S42/T3** samples is that the CD spectra for the cross-linked duplexes continued to display the characteristic B-DNA bands and therefore the overall characteristic structure is not significantly altered by the cross-link. Both Figures 10 and 11 indicate a small decrease in the intensities of both positive and negative bands, which would suggest a small degree of unwinding or unstacking of the double helix structure in order to accommodate the newly formed photodimer. It is probable that some of the nucleobases adjacent to the site of the cross-link are forced off axis by the photodimer, destabilising the interactions with flanking bases and leading to the changes that are seen in the CD spectrum.

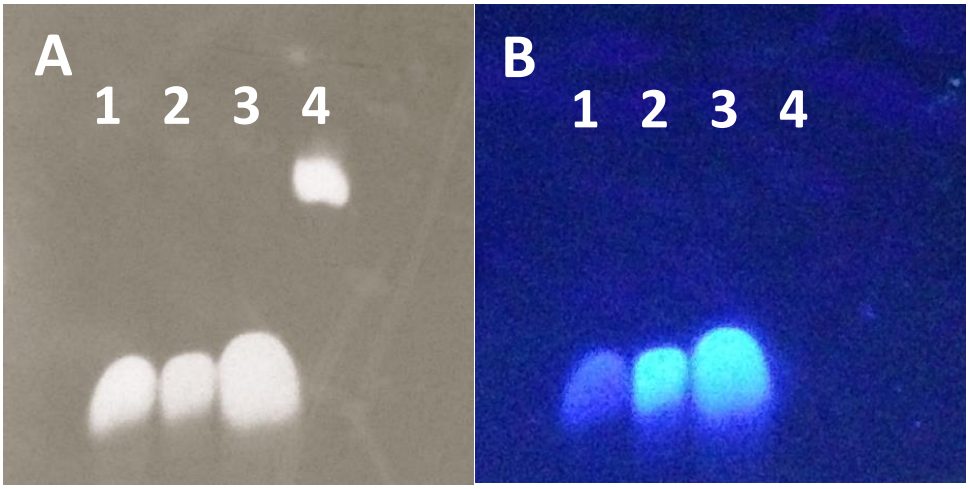
6.3.5 Gel Electrophoresis.

Gel electrophoresis was used to confirm the T_m and CD observations by studying how the modified sequences migrate through a gel before and after photodimerisation. In this study two experiments were set up; native conditions and denaturing conditions, using the same set of samples for each test, (Figures 12 & 13 respectively). Under native conditions the modified cross-linked duplex **S41/T2-CL** was expected to behave in exactly the same way as that of the starting material. CD results indicated that there was possibly some disruption to the B-DNA type structure upon photodimerisation but this was not significant enough to warrant a sizeable difference in migration speed (Figure 13, lanes 3 & 4).



Lane	1	2	3	4
Sample	**S41**	**T2**	**S41/T2**	**S41/T2-CL**

Figure 12. Native gel electrophoresis experiment. The similarity in bands for lanes 3 & 4 confirm that anthracene cross-linking does not significantly disrupt the B-DNA structure. (20% native PAGE gel, 50 mM NaCl, 1 x TB buffer, 20 hour run time, 100V)



Lane	1	2	3	4
Sample	**S41**	**T2**	**S41/T2**	**S41/T2-CL**

*Figure 13. Denaturing gel electrophoresis experiment. The lower migration rate of the cross-linked sample (**S41/T2-CL**) indicates the cross-link has prevented strand dissociation under these conditions. (A) Short wavelength – 254 nm (B) Long wavelength – 365 nm. (20% denaturing PAGE gel, 7 M urea, 1 x TBE buffer, 3 hour run time, 20W)*

Under denaturing conditions, the use of urea within the gel prevents the formation of hydrogen bonds and hence no secondary structure should form. Therefore in normal circumstances all sequences should dissociate and travel through the gel at the same rate as they each have similar masses to each other. As predicted under these denaturing conditions, figure 13A indicated that all samples migrated through the gel with similar migration speeds, with the exception of the cross-linked product in lane 4. With the cross-link in place the ICL product cannot dissociate under the denaturing conditions, and whilst there should not be any base pairing, the sample should still behave as an aggregate and travel at a slower rate due to its larger cross-sectional area. Despite earlier analysis, which found the anthracene ICL product to revert to the starting materials at elevated temperatures, this result is good evidence that the anthracene dimer cross-link is sufficiently robust for it to withstand processes that would normally dissociate a DNA duplex. As a secondary measure the same denaturing gel slab was analysed using a longer

wavelength of UV light; the resulting image will pick up the non-dimerised anthracene groups within the oligonucleotide samples that fluoresce under these UV conditions. In Figure 13B, only the cross-linked sample in lane 4 does not show up as the anthracene photodimer is not fluorescent. This result shows how fluorescent imaging could be used as a rapid validation method to confirm that cross-linking has taken place, which could prove to be useful in biological experiments involving *in-vitro* cell studies.

6.4 Conclusions and Further Work

The cross-linking of complementary sequences through anthracene dimerisation has proven to be a fast, effective and reliable technique and has the potential to be applied to wide range of new studies. The insertion of anthracene tags within modified complimentary oligonucleotides can selectively cross-link to give well-defined photoproducts that are easily characterised and are useful for the study of enzymatic repair pathways. Whilst not as thermally stable as intra-strand photodimers (Chapter 3) the anthracene photodimer that creates the cross-link between complementary strands has proven to be robust enough to withstand dissociative conditions used in denaturing gel electrophoresis. Therefore multiple, selective insertions of anthracene tags into relevant sequences could help reversibly strengthen DNA architectures such as those developed by Seeman or Sleiman (Chapter 1.3).

Many natural biological processes rely on dissociation of the double stranded duplex structure in order to read the genetic information carried within DNA. Therefore control over these processes using light would be an interesting focus of research. One particular

target of research would be the helicase family of proteins, the primary agent responsible for strand separation during transcription, DNA replication and DNA repair.[23] Helicases are donut shaped motor proteins, which thread over single stranded overhangs of a genetic duplex and advance along the strand, using ATP hydrolysis as the driving force (Figure 14). This separation creates a single stranded bubble through which the genetic information from the single stranded components can be read; however a cross-link between the strands could potentially interrupt this process.[24]

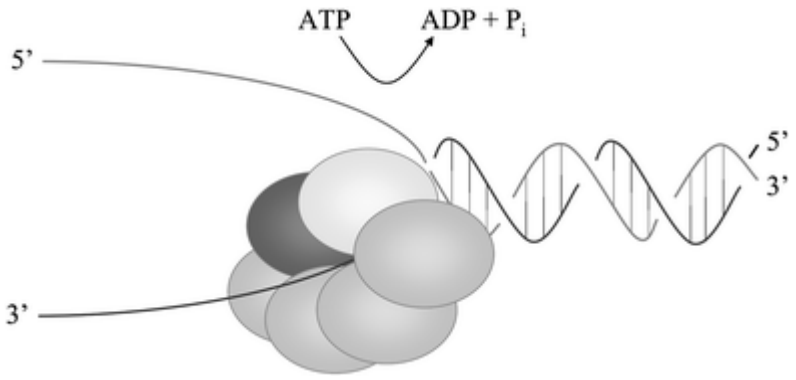
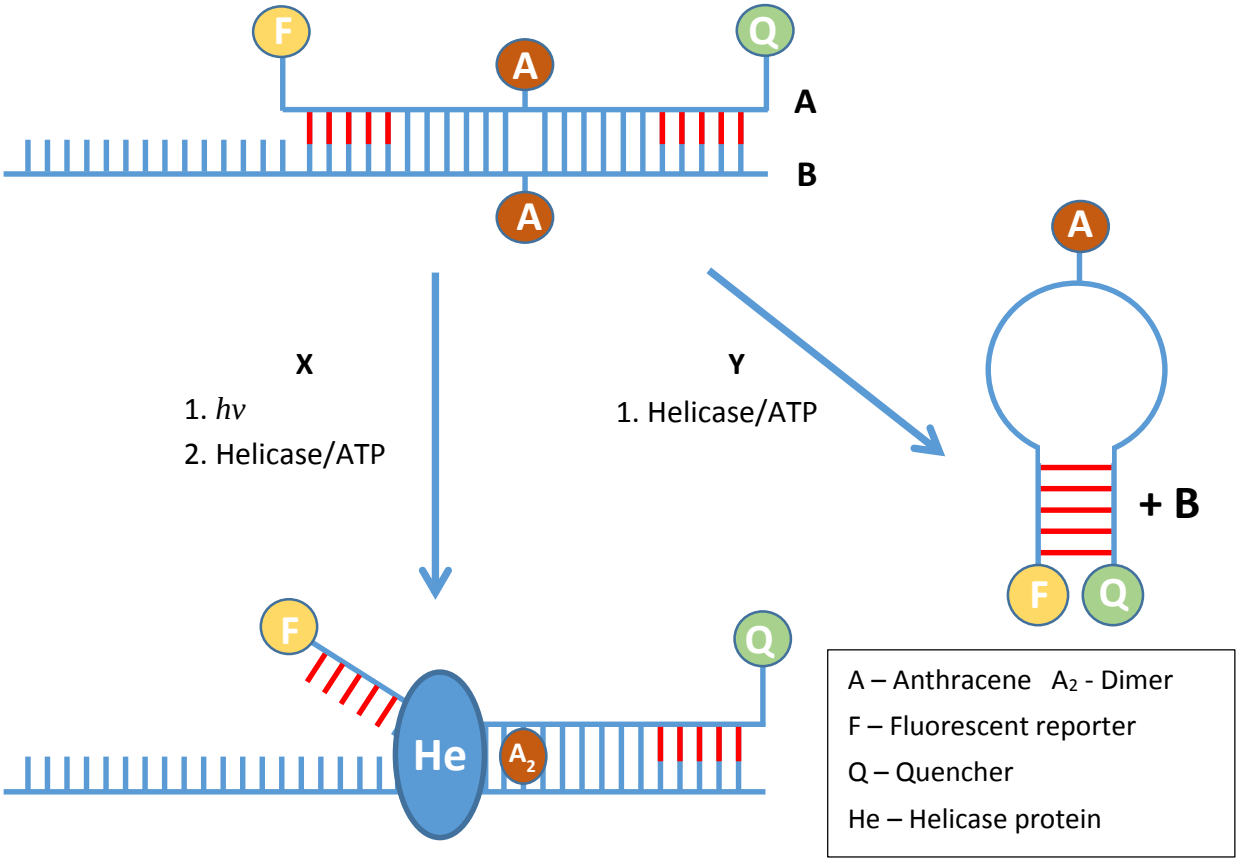


Figure 14. A schematic diagram of a helicase protein as it unwinds a double stranded DNA sequence. The donut shaped protein threads over a single stranded segment of DNA and drives along the sequence with the use of ATP hydrolysis. Since the internal cavity of the protein is only large enough to accommodate single stranded DNA the base pairs are forced to separate. Image taken from ref. 26.

The main issue with this research is that there are few techniques that allow the disruption of *helicase* activity to be monitored without significant drawbacks. However some promising research by Frick and co-workers has discovered a new FRET based assay that can monitor helicase activity in real time.[25] The method involves specifically designed oligonucleotide sequences that incorporate fluorescent reporter and quencher groups, which provide a varying fluorescent response depending on whether the helicase has successfully separated the two component strands. The key concept to this technique is a hairpin trap which prevents the two strands from re-annealing after separation,

which is implemented through the rational design of two component sequences (**A** + **B**, Scheme 6). A helicase protein binds to a short single stranded overhang on strand **B** and moves along towards the opposite end and as it does so dissociates the secondary strand (**A**). Self-complementary segments within strand **A** (in red) then bind with each other to form the hairpin structure, bringing the fluorophore and quencher groups on opposite ends into close proximity and reducing the fluorescent output from the system. It would be interesting to determine whether further modification of this system with anthracene groups within the component strands would make it possible to create a cross-link that could potentially block the progress of the helicase protein as it moves along the



*Scheme 6. Control over helicase activity through anthracene photodimerisation. On the addition of a helicase protein the sequences **A** + **B** are divided, a stem loop is formed and the fluorescent signal is quenched. However the formation of an anthracene cross-link between **A** + **B** blocks helicase movement and the fluorescent signal remains.*

oligonucleotide (Scheme 6). The blocking of the helicase mobility could then be detected through the limited or non-response from the hairpin trap. If this technique was successful it would allow control over helicase activity through light, enabling researchers to better study their role within biological systems.

There are other known methods of cross-linking, which can be utilised in this context, one such example is psoralen. Psoralen is a planar, tricyclic organic compound (Figure 15A) found naturally in a variety of sources that can be used for the treatment of numerous disorders including psoriasis, eczema and vitiligo. The planar aromatic structure of psoralen allows it to intercalate between the bases within DNA, much like anthracene, and upon absorption of UV light the psoralen undergoes consecutive 2 + 2 cycloaddition reactions to create an interstrand cross-link, (Figure 15B).[27] This property has been exploited for the treatment of psoriasis due to the fact that the induced ICL acts as an anti-proliferative and inhibits the rapid cell growth that initiate the symptoms of the condition. It would be interesting to see if an anthracene initiated ICL could mirror this anti-proliferative effect and investigate whether the reversible nature of the anthracene cross-link could play a role in the study of genetic disorders such as psoriasis, the causes of which are not fully understood.

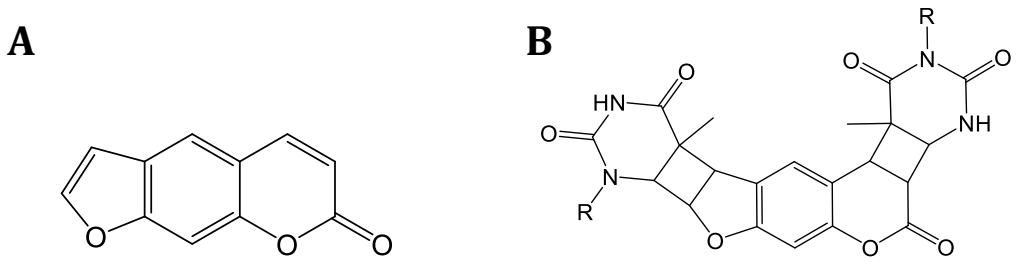


Figure 15: (A) Structure of the ICL forming compound psoralen. (B) Following consecutive 2 + 2 cycloaddition reactions psoralen creates an interstrand cross-link between two neighbouring thymine nucleobases.

6.5 References

1. Jackson, S. P.; Bartek J., *Nature*, **2009**, *461*, 1071-1078.
2. Jackson, A. L.; Loeb, L. A., *Mutation Research*, **2001**, *477*, 7-21.
3. Poirier, M. C., *Nature Reviews Cancer*, **2004**, *4*, 630-637.
4. Sinha, R. P.; Häder, D-P., *Photochem. Photobiol. Sci.*, **2002**, *1*, 225-236.
5. Auerbach, C.; Robson, J. M.; Carr, J. G., *Science*, **1947**, *105*, 243-247.
6. Mouret, S.; Baudouin, C.; Charveron, M.; Favier, A.; Cadet, J.; Douki, T., *Proc. Natl. Acad. Sci. USA*, **2006**, *103*, 13765-13770.
7. Brash, D. E.; Rudolph, J. A.; Simon, J. A.; Lin, A.; McKenna, G. J.; Baden, H. P.; Halperin, A. J.; Pontén, J., *Proc. Natl. Acad. Sci. USA*, **1991**, *88*, 10124-10128.
8. Noll, D. M.; Mason, T. M.; Miller, P. S., *Chem. Rev.*, **2006**, *106*, 277-301.
9. Deans, A. J.; West, S. C.; *Nature Reviews Cancer*, **2011**, *11*, 467-480.
10. Schärer, O. D., *ChemBioChem.*, **2005**, *6*, 27-32.
11. Rajski, S. R.; Williams, R. M., *Chem. Rev.*, **1998**, *98*, 2723-2795.
12. Poklar, N.; Pilch, D. S.; Lippard, S. J.; Redding, E. A.; Dunham, S. U.; Breslauer, K. J., *Proc. Natl. Acad. Sci. USA*, **1996**, *93*, 7606-7611.
13. Li, X.; Heyer, W-D., *Cell Research*, **2008**, *18*, 99-113.
14. M., Op de Beeck; Madder, A., *J. Am. Chem. Soc.*, **2011**, *133*, 796-807.
15. Stevens, K.; Madder, A., *Nucleic Acids Res.,* **2009**, *37*, 1555-1565.
16. Stevens, K.; Madder, A., *Nucleosides, Nucleotides, Nucleic Acids,* **2007**, *26*, 1359-1362.
17. Kashida, H.; Doi, T.; Sakakibara, T.; Hayashi, T.; Asanuma, H., *J. Am. Chem. Soc.*, **2013**, *135*, 7960-7966.
18. Lewis, F.D.; Wu, T.; Burch E. L.; Bassani, D. M.; Yang, J.-S.; Schneider, S.; Jäger, W.; Letsinger, R. L., *J. Am. Chem. Soc.*, **1995**, *117*, 8785-8792.
19. Tagawa, M.; Shohda, K.; Fujimoto, K.; Sugawara, T.; Suyama, A., *Nucleic Acids Res.*, **2007**, *35*, e140.
20. Shelbourne, M.; Brown, T.; El-Sagheer, A. F.; Brown, T., *Chem. Commun.*, **2012**, *48*, 11184-11186.
21. Shelbourne, M.; Chen, X.; Brown, T.; El-Sagheer, A. F., *Chem. Commun.*, **2011**, *47*, 6275-6259.
22. Schildkraut, C.; Lifson, S., *Biopolymers*, **1965**, *3*, 195-208.
23. Umate, P.; Tuteja, N.; Tuteja, R., *Commun. Integr. Biol.*, **2011**, *4*, 118-137.
24. Cortez, D., *Genes & Dev.*, **2005**, *19*, 1007-1012.
25. Belon, C. A.; Frick, D. N., *BioTechniques*, **2008**, *45*, 433-442.
26. Vindigni, A., *Mol. BioSyst.*, **2007**, *3*, 266-274.
27. Bethea, D.; Fullmer, B.; Syed, S.; Seltzer, G.; Tiano, J.; Rischko, C.; Gillespie, L.; Brown, D.; Gasparro, F. P., *J. Dermatol. Sci.*, **1999**, *19*, 78-88.

Chapter 7: Experimental

7.1 Synthesis and Experimental Details

7.1.1 Reagents and Chemicals

Solvents and reagents were obtained from commercial suppliers and used without further purification. Anhydrous solvents were prepared using an IT PureSolv purification system or otherwise prepared by the usual procedures. All reactions were carried out in the absence of light, under an argon atmosphere due to the sensitivity of anthracene. Column chromatography was carried out using Silica Gel (Merck, grade 60) or alumina (basic, Brockman activity I).

^{1}H, ^{13}C and ^{31}P NMR spectra were recorded on a Bruker AVIII300 or AVIII400. ES^- mass spectra were measured by a Waters Micromass LCT time of flight mass spectrometer. Anthracene monomers utilising the n = 6 linker were prepared from a stock of the DMT protected compound synthesised by J.L. Duprey. Compounds JLD1 and JLD2 were tested for purity and compared to published assignments before use.

7.1.2 DNA Synthesis

Solid phase, automated DNA synthesis was carried out at the University of Birmingham using an Applied Biosciences ABI 394 DNA/RNA synthesiser on a 1 mM scale. All phosphoramidites, reagents and columns were purchased from Link Technologies. Oligonucleotide synthesis grade acetonitrile and trichloroacetic acid was purchased from

Rathburn chemicals and Sigma Aldrich respectively. For use in the stem loop experiments Cy3 and Cy5 dyes were purchased as phosphoramidites from Glenn Research and incorporated into the oligonucleotides during automated synthesis. In order to avoid degradation ultra-mild reagents were used to synthesise Cy3 or Cy5 labelled oligonucleotides, which was also purchased from Glenn Research. HPLC or DNA grade water was used for the synthesis and purification of all oligonucleotides synthesised.

Oligonucleotides bearing a 3' anthracene modification were synthesised using a 3'-phosphate column to add the anthracene as the first monomer within the sequence. Therefore these sequences terminate with a 3'-phosphate rather than the normal 3'-hydroxyl group.

7.1.3 DNA Purification through HPLC

Preparative HPLC purification was carried out using a Dionex Summit P680 pump coupled with a Summit UVD1705 detector fitted with a prep flow cell. A Phenomenex Clarity 5µ Oligo RP column 150 x 10 mm column was used for the purification. Following HPLC isolation the desired product underwent a final desalting workup using an illustra NAP-10 gel filtration column and stored in DNA grade water using a UV-blocking container.

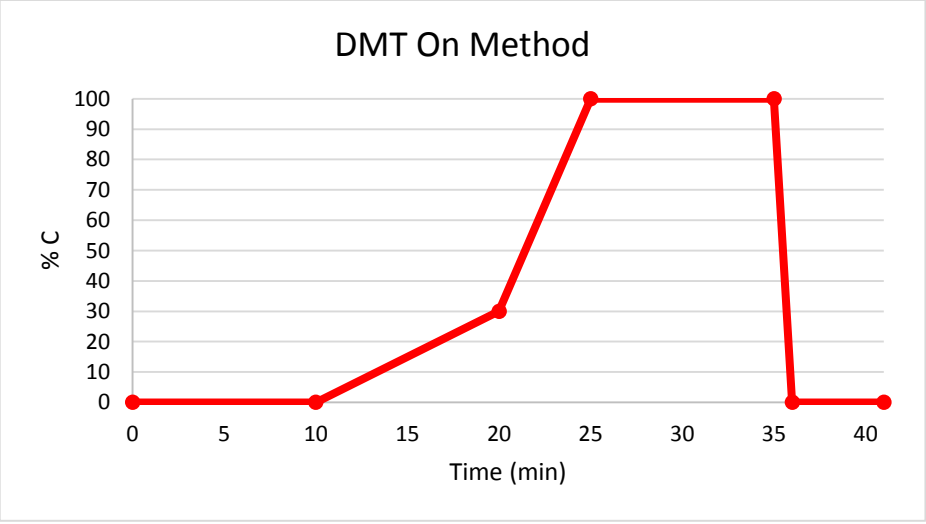
Two HPLC methods were used for the purification and analysis of oligonucleotides in this project depending on the presence of an anthracene modification. An anthracene modification acts much like a 5'-DMT and increases the hydrophobicity of the full length oligonucleotide making it easier to purify from other truncated components. Any

sequence bearing a modification were therefore purified using the 'DMT-On' method, which is shorter and uses a higher percentage of acetonitrile. Control and complementary sequences not bearing an anthracene modification were purified using the 'Oligo SP' method, which is longer and utilises a more gradual increase in the percentage of acetonitrile.

Solvents:

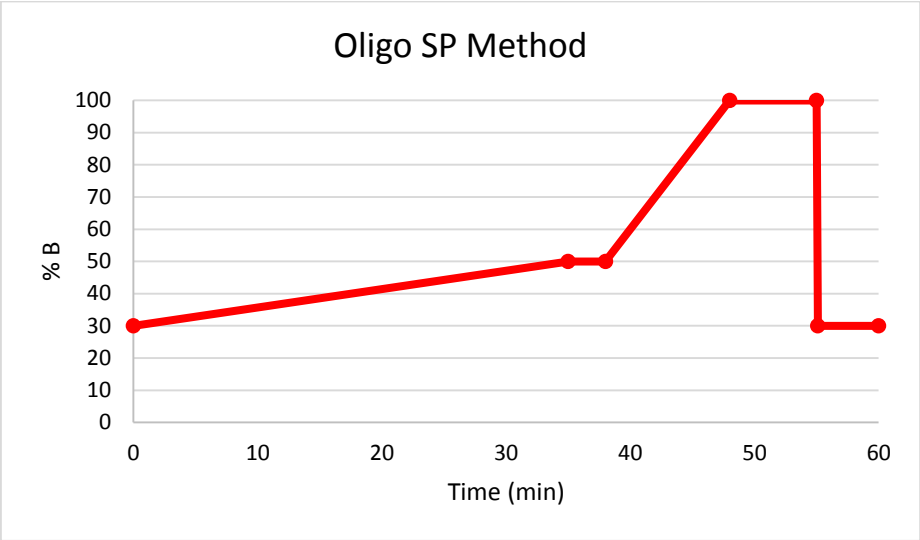
A. 0.1M TEAA (pH 7.0), 5% CH_3CN
B. 0.1M TEAA (pH 7.0), 15% CH_3CN
C. CH_3CN

<u>DMT On method:</u>



0-10 mins, 100% B; 10-20 mins, 0% C – 30% C (remainder B); 20-25 mins, 30% C – 100% C; 25-35 mins, 100% C; 35-36 mins, 100% C – 0% C; 36-41 mins, 100% B.

Oligo SP method:



0-35 mins, 30% B – 50% B (remainder A); 35-38 mins, 50% B; 38-48 mins, 50% - 100% B; 48-55 mins, 100% B; 55-55.1 mins, 100% B – 30% B; 55.1-60 mins, 30% B.

Following purification of the oligonucleotides through preparative HPLC all samples were stored in DNA grade water and separated into a number of aliquots for freezer storage.

7.1.4 Oligonucleotide Characterisation

All oligonucleotides were characterised by analytical HPLC, which was carried out using a Shimadzu Prominence UFLC kit fitted with an auto-sampler. A Phenomenex Clarity 5µ Oligo RP column 150 x 4.6 mm was used for the HPLC analysis and the same methods were used as discussed previously. Further mass spec characterisation was obtained using a Waters Micromass LCT mass spectrometer, which uses maximum entropy software to generate a full mass spectrum from the raw data. In order to do this the

software analyses the isotopic distribution of each peak to deconvolute the spectrum data and generate a molecular mass of the neutral species.

All oligonucleotide concentrations were measured through UV spectroscopy using either a Varian Cary 50 or Shimadzu 1800 UV spectrometer. The values were determined through the Beer-Lambert law where the value for ε is calculated by online software.

7.1.5 Photochemistry

Anthracene photodimerisation studies were performed by making up a 2 mL solution containing 20 μM of the sample oligonucleotide, with 10 mM sodium phosphate buffer and 100 mM NaCl. Each sample was degassed with argon for at least ten minutes before testing to help remove oxygen. Samples were irradiated using a Photochemical Reactors Ltd 125W mercury arc lamp, which is placed in a housing that has a small outlet where a 365 nm bandpass filter is fitted, (Edmund Optics 65-615, 12.5 mm diameter, 10 nm bandwidth). Each sample was irradiated with continuous degassing, which not only prevents formation of anthraquinone but also aids stirring of the reaction mixture. The reaction was monitored at set intervals through UV spectroscopy by measuring the change in absorbance of each sample at 260 and 365 nm, the reaction was stopped when no further change in the absorbance was observed.

7.1.6 UV/Vis and CD Spectroscopy

UV/Vis spectra were recorded using either a Varian Cary 50 or Shimadzu 1800 Spectrometer. Variable Temperature UV experiments were carried out on a Varian Cary 5000 Spectrometer equipped with a peltier and a six sample auto-changer.

DNA T_m Melting Point Experiments

To measure the melting curve of a sample duplex a 500 μL solution was prepared containing 5 μM of the sample oligonucleotide and its complementary strand along with 10 mM sodium phosphate buffer and 100 mM NaCl. The sample was analysed through three successive temperature ramps, two dissociative and one annealing between 15 and 85 °C (unless otherwise stated), recording the absorbance at 260 nm at a rate of 0.5 °C per minute. The final T_m value is determined by a derivative calculation and the reported value is an average of the two dissociative ramps.

For quadruplex samples the T_m is calculated from measuring the hypochromic shift in UV absorbance at 297 nm, however all other parameters for the measurement are kept the same. Quadruplex samples are prepared to a 5μM concentration using 50 mM KCl and 100 mM Tris.HCl pH 7.0 buffer to ensure quadruplex folding.

For the anthracene photodimers the temperature range was in some cases restricted to 15-55 °C to prevent thermal dissociation of the anthracene dimer, which can introduce errors within the data. All T_m data collected using this temperature range were checked to ensure that the entire melting/annealing transition was complete within the restricted

range so as not to introduce errors when the final T_m value is calculated. In other cases the range was lowered to 5 °C in order to observe any low temperature T_m transitions, in this situation a nitrogen supply was fed into the sample chamber to prevent any condensation build up on the sample cuvettes.

Circular Dichroism Spectroscopy

Circular Dichroism spectra were recorded on a Jasco J-810 spectropolarimeter between 200 and 450 nm with a scanning rate of 100 nm/min, at medium sensitivity. The final spectra is recorded with 10 accumulations. A baseline scan is measured separately and subtracted from the data after scanning.

7.1.7 Gel Electrophoresis

All gel electrophoresis experiments were carried out using a Hoefer SE400 vertical slab gel kit paired with a Biorad PowerPac HV 5000V/500mA/400W power supply. The gel electrophoresis kit allows for 24 cm gels to be cast with a 0.5-1 mm thickness and a maximum of 15 sample wells. Below are listed the general conditions used for each of the gel electrophoresis experiments performed within this project:

Native Gel Electrophoresis

For native conditions a 20 % poly acrylamide gel was used containing 50 mM NaCl and 1 x TB buffer, 1 OD of the oligonucleotide samples were loaded onto the gel and run at 100V

for approximately 16 to 20 hours. Note: for the TBA/thrombin gels, KCl was used instead of NaCl. All samples were heated and allowed to cool before loading to aid annealing for samples that contain complementary sequences.

Denaturing conditions

For denaturing conditions a 20% polyacrylamide gel was used containing 7M urea and 1 x TBE buffer, 1 OD of the oligonucleotide samples were loaded onto the gel and run at 20W for approximately 2 to 3 hours.

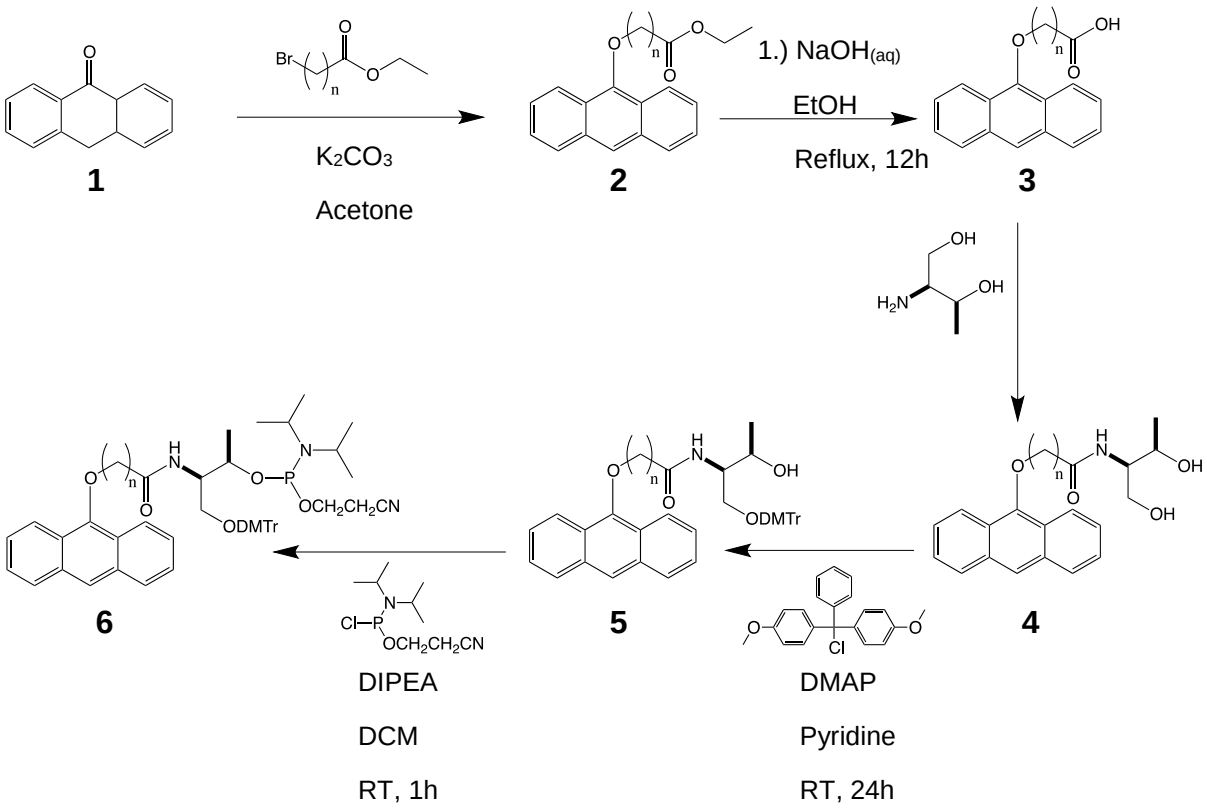
All gel electrophoresis experiments were tracked using the standard dyes bromophenol blue and xylene cyanol. The gel slabs were imaged using the UV shadow technique and if required were stained with either ethidium bromide or SYBR gold. Imaging of the gel stains were performed using an Alpha Innotech AlphaImager HP imaging system.

7.1.8 Fluorescence

Fluorescence titrations were carried out on a Shimadzu RF-5301PC spectrofluorimeter, using a 365 nm excitation wavelength and the fluorescence recorded between 370 and 570 nm. Samples were measured at a 1 μM concentration under standard conditions (100 mM NaCl, 10 mM phosphate buffer, ambient temperature).

7.2 Synthesis of Anthracene Monomers

7.2.1 Reaction Scheme



Scheme 1. Synthesis route for the production of n = 4, 6 and 7 anthracene monomers starting from the commercially available anthrone.

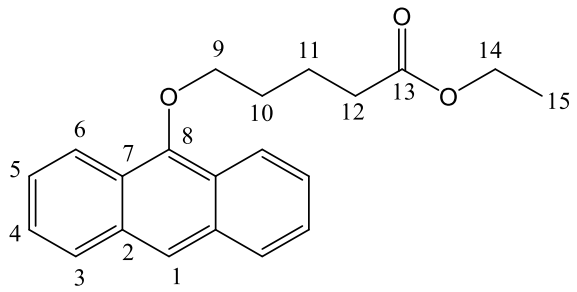
The following sections describe the synthesis of the various anthracene phosphoramidites produced for this project according to the scheme above. Each section describes the synthesis of different variations of each compound depending on the type of brominated ester (*n* = 4,6,7) and isomer of threoninol (D/L) used to produce the linker group.

7.2.2 Synthesis of Anthracene Ester **2**:

Anthrone (**1**) and K_2CO_3 were dissolved in degassed acetone (200 mL) and stirred under N_2 at room temperature in the absence of light for 15 min. A brominated ester was added to the reaction mixture and refluxed overnight. After reaction completion the solution was cooled to room temperature and filtered to remove the solid K_2CO_3. The solution was then evaporated in *vacuo* and re-dissolved in DCM (100 mL). The solution was then washed with H_2O (1 x 50 mL) and dried over $MgSO_4$. The solvent was removed in *vacuo* to give an orange oil. The final product was isolated through column chromatography on silica (hexane with 10% ethyl acetate) to give a white creamy solid.

Ethyl 5-(anthracen-9-yloxy)pentanoate – (*n* = 4)

(JM1.1)



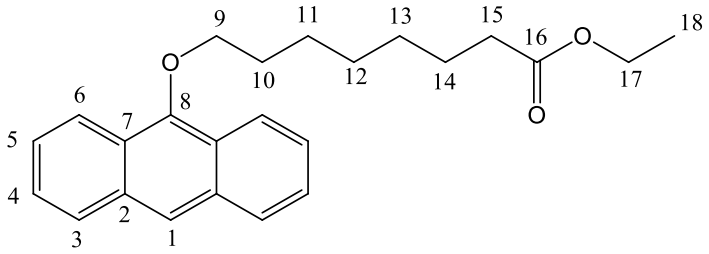
Anthrone	5.832 g	30 mM
K_2CO_3	4.154 g	30 mM
ethyl 5-bromopentanoate	4.8 mL	30 mM
Yield	4.4493 g	46.06 %

M.p. 70 ˚C; δ_H (300 MHz, $CDCl_3$) 8.30 (2H, d, J = 9.1, H_6), 8.24 (1H, s, H_1), 7.93 (2H, d, J = 9.7, H_3), 7.44-7.55 (4H, m, H_5 & H_4), 4.17-4.26 (4H, m, H_{14} & H_9), 2.53 (2H, t, J = 7.5, H_{12}), 2.00-2.12 (4H, m, H_{11} & H_{10}), 1.31 (3H, t, J = 7.1, H_{15}); δ_C (400 MHz, $CDCl_3$) 173.8 (C_{13}), 151.1 (C_8), 132.2 (2xC_2), 128.4 (2xC_3), 125.4 (2xC_5), 125.1 (2xC_4), 124.6 (2xC_7), 122.2

(2xC_6), 122.1 (C_1), 75.5 (C_9), 60.4 (C_{14}), 58.5 (C_{10}), 34.3 (C_{12}), 30.1 (C_{11}), 14.3 (C_{15}); *m/z* (ES+) calcd. For $C_{21}H_{22}O_3Na$ ($M+Na^+$) 355.2 found 355.2.

Ethyl 8-(anthracen-9-yloxy)octanoate – (*n* = 7)

(JM5.1)



Anthrone	3.86 g	19.9 mM
K_2CO_3	5 g	36.2 mM
ethyl 8-bromooctanoate	4.2 mL	19.9 mM
Yield	4.3044 g	59.42 %

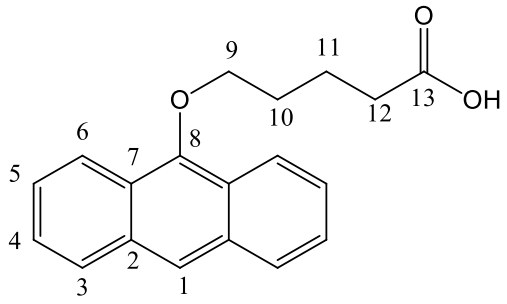
(R_f = 0.62 in hexane with 10% EtOAc); δ_H (300 MHz, $CDCl_3$) 8.28 (2H, d, *J* = 9.1, H_6), 8.21 (1H, s, H_1), 7.98 (2H, d, *J* = 10.6, H_3), 7.45-7.47 (4H, m, H_4 & H_5), 4.10-4.20 (4H, m, H_9 & H_{17}), 2.33 (2H, t, *J* = 7.56, H_{15}), 2.05 (2H, quintet, *J* = 6.63, H_{10}), 1.63-1.72 (4H, m, H_{11} & H_{14}), 1.42-1.50 (4H, m, H_{12} & H_{13}), 1.26 (3H, t, *J* = 7.14, H_{18}); δ_C (400 MHz, $CDCl_3$) 173.8 (C_{16}), 151.5 (C_8), 132.4 (2xC_2), 128.4 (2xC_3), 125.4 (2xC_4), 125.0 (2xC_5), 124.7 (2xC_7), 122.4 (C_1), 121.9 (2xC_6), 76.1 (C_9), 60.2 (C_{17}), 34.4 (C_{15}), 30.6 (C_{10}), 29.3 (C_{12}), 29.1 (C_{13}), 26.1 (C_{11}), 25.0, (C_{14}), 14.3 (C_{18}). *m/z* (ES+) calcd. for $C_{24}H_{28}O_3Na$ ($M+Na^+$) 387.19, found 387.19.

7.2.3 Synthesis of Carboxylic Acid **3**

The anthracene ester **2** is dissolved in a 1:1 solution of ethanol and 10% NaOH (a.q.) (300 mL) and refluxed overnight under a N_2 atmosphere in the absence of light. On completion the solution was concentrated in *vacuo* and dissolved in H_2O (500 mL). Concentrated HCl was then added drop wise to the solution until a precipitate formed, which was collected by suction filtration. The solid was washed with water and dried to give the desired product as a white solid.

5-(Anthracen-9-yloxy)pentanoic acid – (*n* = 4)

(JM1.2)

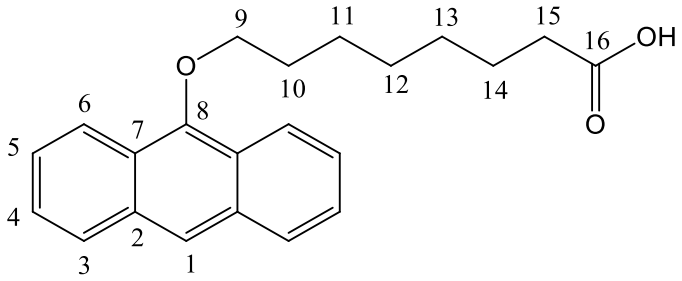


JM1.1	4.03 g	12.5 mM
Yield	3.3927	92.2 %

M.p. 117 ˚C; δ_H (300 MHz, $CDCl_3$) 8.29 (2H, d, *J* = 9.3, H_6), 8.24 (1H, s, H_1), 7.99 (2H, d, *J* = 9.4, H_3), 7.44-7.55 (4H, m, H_5 & H_4), 4.25 (2H, t, *J* = 5.9, H_9), 2.62 (2H, t, *J* = 6.9, H_{12}), 2.10-2.16 (4H, m, H_{11}, & H_{10}); δ_C (75 MHz, $CDCl_3$) 173.8 (C_{13}), 151.1 (C_8), 132.2 (2xC_2), 128.4 (2xC_3), 125.4 (2xC_5), 125.1 (2xC_4), 124.6 (2xC_7), 122.2 (2xC_6), 122.1 (C_1), 75.5 (C_9), 33.8 (C_{12}), 30.0 (C_{10}), 21.6 (C_{11}); *m/z* (ES+) calcd. For $C_{19}H_{18}O_3Na$ ($M+Na^+$) 317.1 found 317.1.

8-(Anthracen-9-yloxy)octanoic acid – (*n* = 7)

(JM5.2)



JM5.1	4.3044 g	11.8 mM
Yield	2.686 g	67.6 %

(R_f = 0.61 in DCM with 5% MeOH); M.p. 107-109 °C; ν_{max}/cm^{-1} 1709 (C=O); δ_H (300 MHz, $CDCl_3$) 8.28 (2H, d, *J* = 9, H_6), 8.21 (1H, S, H_1), 7.98 (2H, d, *J* = 7.99, H_3), 7.43-7.50 (4H, m, H_4 & H_5), 4.19 (2H, t, *J* = 6.63, H_9), 2.41 (2H, t, *J* = 7.45, H_{15}), 2.04 (2H, quintet, *J* = 6.6, H_{10}), 1.63-1.77 (4H, m, H_{11} & H_{14}), 1.43-1.52 (4H, m, H_{12} & H_{13}); δ_C (400 MHz, $CDCl_3$) 178.4 (C_{16}), 155.3 (C_8), 132.5 (2xC_2), 128.4 (2xC_3), 125.4 (2xC_4), 125.0 (2xC_5), 124.7 (2xC_7), 122.4 (C_1), 121.9 (2xC_6), 76.1 (C_9), 33.8 (C_{15}), 30.6 (C_{10}), 29.2 (C_{12}), 29.1 (C_{13}), 26.1 (C_{11}), 24.7 (C_{14}). *m/z* (ES+) calcd. for $C_{22}H_{23}O_3$ (M+) 335.1725, found 335.1648.

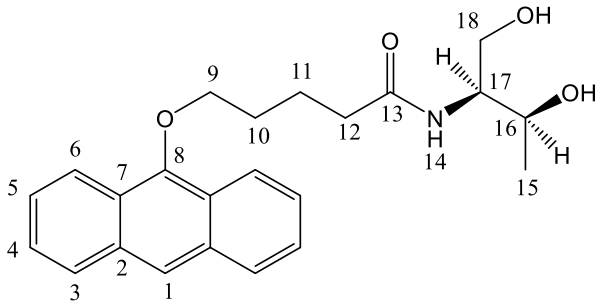
7.2.4 Synthesis of Anthracene Diol **4**

The carboxylic acid product **3** and HBTU was dissolved in anhydrous DMF (20 mL) and stirred under N_2 at room temperature in the absence of light for 15 min. Threoninol (D/L) and DIPEA were added to the solution and the reaction was stirred at 40 °C for 40 hours. The solution was mixed with 1:2 methanol/DCM (100 mL), washed thoroughly with H_2O (6 x 50 mL) and dried over $MgSO_4$. The solvent was removed under *vacuo* and the final

product was isolated through column chromatography on silica (DCM with 5% methanol) to give an oily yellow solid.

5-(Anthracen-9-yloxy)-N-((2*S*,3*S*)-1,3-dihydroxybutan-2-yl)pentanamide – (*n* = 4D)

(JM1.3D)

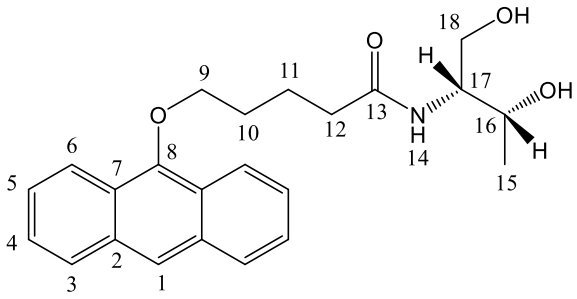


JM1.2	0.5 g	1.7 mM
HBTU	0.64 g	1.7 mM
D-Threoninol	0.2 g	1.9 mM
DIPEA	0.4 mL	2.3 mM
Yield	299.4 mg	46 %

M.p. 126-131 ˚C; δ_H (400 MHz, $CDCl_3$) 8.24 (2H, dd, *J* = 7.9 & 5.8, H_6), 8.18 (1H, s, H_1), 7.90-7.98 (2H, m, H_3), 7.42-7.51 (4H, m, H_5 & H_4), 6.57 (1H, d, *J* = 8.5, N*H*), 4.05-4.19 (3H, m, H_9 and H_{16}), 3.86 (1H, ddd, *J* = 16.9, 9.1 & 6.6, H_{17}), 3.72-3.81 (2H, m, H_{18}),2.41 (2H, t, *J* = 6.6, H_{12}), 2.06 (4H, m, H_{11} & H_{10}), 1.18 (3H, d, *J* = 6.4, H_{15}); δ_C (100 MHz, $CDCl_3$) 173.8 (C_{13}), 151.1 (C_8), 132.2 (2xC_2), 128.4 (2xC_3), 125.4 (2xC_5), 125.1 (2xC_4), 124.6 (2xC_7), 122.2 (2xC_6), 122.1 (C_1), 75.5 (C_9), 68.2 (C_{16}),64.5 (C_{18}), 54.8 (C_{17}), 36.5 (C_{12}), 30.1 (C_{10}), 22.7 (C_{11}), 20.4 (C_{15}); *m/z* (ES+) calcd for $C_{23}H_{27}O_4NNa$ ($M+Na^+$) 404.18, found 404.18.

5-(Anthracen-9-yloxy)-N-((2R,3R)-1,3-dihydroxybutan-2-yl)pentanamide – (*n* = 4L)

(JM1.3L)

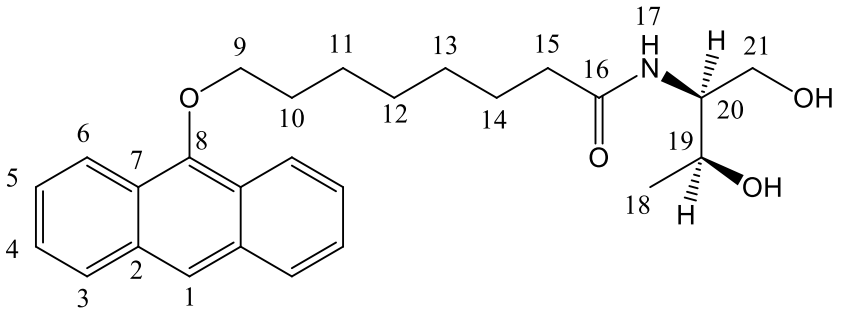


JM1.2	0.5 g	1.7 mM
HBTU	0.64g	1.7 mM
L-Threoninol	0.2 g	1.9 mM
DIPEA	0.4 mL	2.3 mM
Yield	401.7 mg	62 %

M.p. 126-131 ˚C; δ_H (400 MHz, $CDCl_3$) 8.24 (2H, dd, *J* = 7.9 & 5.8, H_6), 8.18 (1H, s, H_1), 7.90-7.98 (2H, m, H_3), 7.42-7.51 (4H, m, H_5 & H_4), 6.57 (1H, d, *J* = 8.5, H_{14}), 4.05-4.19 (3H, m, H_9 and H_{16}), 3.86 (1H, ddd, *J*= 16.9, 9.1 & 6.6, H_{17}), 3.72-3.81 (2H, m, H_{18}), 2.41 (2H, t, *J* = 6.6, H_{12}), 2.06 (4H, m, H_{11} and H_{10}), 1.18 (3H, d, *J*= 6.4, H_{15}); δ_C (100 MHz, $CDCl_3$) 173.8 (C_{13}), 151.1 (C_8), 132.2 (2xC_2), 128.4 (2xC_3), 125.4 (2xC_5), 125.1 (2xC_4), 124.6 (2xC_7), 122.2 (2xC_6), 122.1 (C_1), 75.5 (C_9), 68.2 (C_{16}), 64.5 (C_{18}), 54.8 (C_{17}), 36.5 (C_{12}), 30.1 (C_{10}), 22.7 (C_{11}), 20.4 (C_{15}); *m/z* (ES+) calcd for $C_{23}H_{27}O_4NNa$ ($M+Na^+$) 404.18, found 404.18.

(2S,3S)-1,3-Dihydroxybutan-2-yl 8-(anthracen-9-yloxy)octanoate – (*n* = 7D)

(JM5.3D)



JM5.2	1.0 g	2.97 mM
HBTU	1.128 g	2.97 mM
D-Threoninol	0.312 g	2.97 mM
DIPEA	0.8 mL	4.6 mM
Yield	0.9103 g	72.3 %

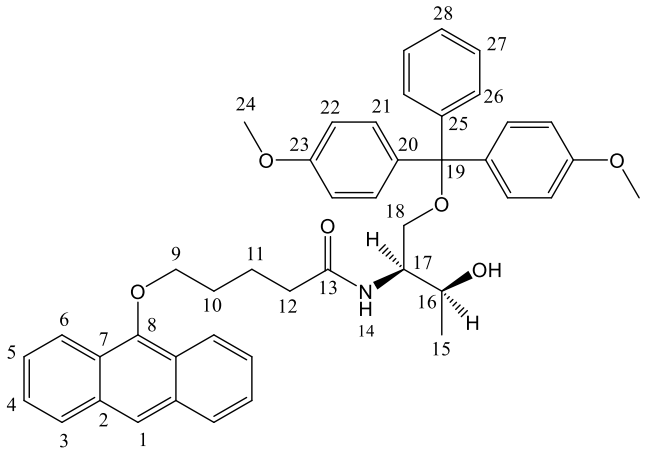
(R_f = 0.47 in DCM with 5% MeOH) ν_{max}/cm^{-1} 3313 (OH), 2929 (N-H), 1674 (C=O); δ_H (300 MHz, $CDCl_3$) 8.26 (2H, d, *J* = 9.6, H_6), 8.19 (1H, s, H_1), 7.96 (2H, d, *J* = 9.47, H_3), 7.43-7.47 (4H, m, H_4 & H_5), 6.35 (1H, d, *J* = 8.23, H_{17}), 4.15 (3H, t, *J* = 6.61, H_9 & H_{19}), 3.74-3.83 (3H, m, H_{20} & H_{21}), 3.54 (1H, t, *J* = 5.01, CHO*H*), 3.43 (1H, d, *J* = 3.39, CH_2OH), 2.25 (2H, t, *J* = 7.57, H_{15}), 2.02 (2H, quintet, *J* = 6.9, H_{10}), 1.59-1.73 (4H, m, H_{11} & H_{14}), 1.38-1.49 (4H, m, H_{12} & H_{13}), 1.17 (3H, d, *J* = 6.37, H_{18}); δ_C (400 MHz, $CDCl_3$) 175.0 (C_{16}), 151.4 (C_8), 132.4 (2xC_2), 128.5 (2xC_3), 125.5 (2xC_4), 125.1 (2xC_5), 124.7 (2xC_7), 122.4 (C_1), 122.0 (2xC_6), 76.1 (C_9), 67.4 (C_{19}), 63.7 (C_{21}), 55.2 (C_{20}), 36.7 (C_{15}), 30.6 (C_{10}), 29.3 (C_{12}), 29.2 (C_{13}), 26.1 (C_{11}), 25.9 (C_{14}), 20.3 (C_{18}). *m/z* (ES+) calcd. for $C_{26}H_{34}NO_4$ (M+) 424.2410, found 424.2487.

7.2.5 Synthesis of DMT Protected Product **5**

The diol product **4**, dimethoxytritylchloride (DMTr) and DMAP were dissolved in anhydrous pyridine (20 mL) and the reaction was left to stir under a N_2 atmosphere at room temperature for 24 hours. On completion the reaction mixture was poured onto H_2O (50 mL) and the product was extracted with DCM (2 x 50 mL). The resulting solution was dried over $MgSO_4$ and the solvent was removed in *vacuo*. The final product was isolated through column chromatography on silica (hexane/ethyl acetate/TEA, 40:59:1) to give a fluffy, yellow, crystalline solid.

5-(Anthracen-9-yloxy)-N-((2S,3S)-1-(bis(4-methoxyphenyl)(phenyl)methoxy)-3-hydroxybutan-2-yl)pentanamide – (*n* = 4D)

(JM1.4D)

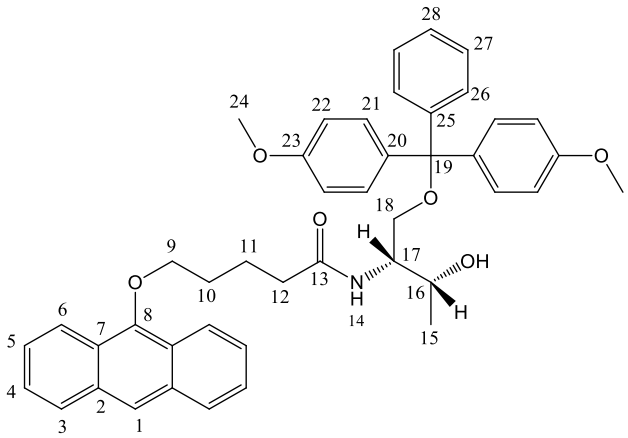


JM1.3D	1.07 g	2.8 mM
DMTr	1.02 g	3 mM
DMAP	0.05 g	0.41 mM
Yield	0.4539 g	23.4 %

M.p. 77-79 ˚C; δ_H (400 MHz, CD_3CN) 8.28 (1H, s, H_1), 8.23-8.27 (2H, m, H_6), 8.03 (2H, dd, J = 6.5 & 2.9, H_3), 7.40-7.49 (6H, m, H_5, H_4 & H_{26}), 7.26-7.31 (6H, m, H_{27} & H_{21}), 7.18 (1H, t, J = 7.3, H_{28}), 6.82 (4H, d, J = 8.8, H_{22}), 6.48 (1H, d, J = 8.8, H_{14}), 4.16 (2H, t, J = 6.1, H_9), 3.90-3.99 (2H, m, H_{16} & H_{17}), 3.69 (6H, s, H_{24}), 3.05-3.19 (3H, m, H_{18} & CHOH), 2.37 (2H, t, J = 7.1, H_{12}), 1.93-2.07 (4H, m, H_{11} and H_{10}), 1.05 (3H, d, J = 6.2, H_{15}); δ_C (75 MHz, CD_3CN) 173.8 (C_{13}), 159.5 (2xC_{23}), 152.2 (C_8), 146.1 (C_{25}), 137.0 (2xC_{20}), 133.4 (2xC_2), 130.9 (4xC_{21}), 129.3 (2xC_3), 129.0 (2xC_{26}), 128.7 (2xC_{27}), 127.7 (C_{28}), 126.6 (2xC_5), 126.3 (2xC_4), 125.5 (2xC_7), 123.1 (2xC_6), 122.9 (C_1), 113.9 (4xC_{22}), 86.8 (C_{19}), 76.6 (C_9), 67.2 (C_{16}), 64.6 (C_{18}), 55.7 (2xC_{24}), 55.2 (C_{17}), 36.7 (C_{12}), 30.8 (C_{10}), 23.4 (C_{11}), 20.5 (C_{15}); m/z (ES+) calcd for $C_{44}H_{45}NO_6$ ($M+Na^+$) 706.31, found 706.31.

5-(Anthracen-9-yloxy)-N-((2R,3R)-1-(bis(4-methoxyphenyl)(phenyl)methoxy)-3-hydroxybutan-2-yl)pentanamide – (n = 4L)

(JM1.4L)

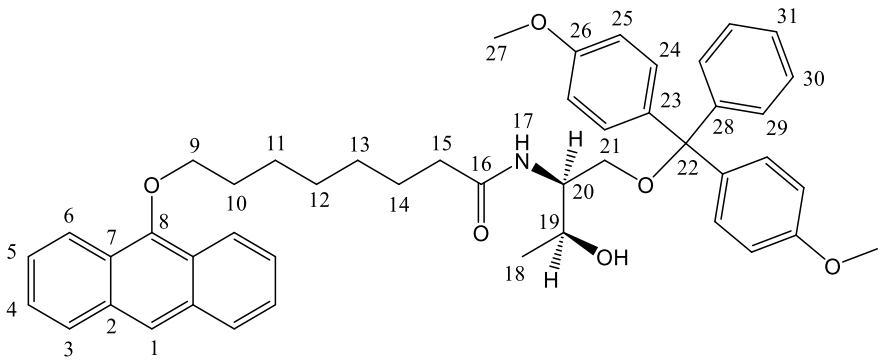


JM1.3L	1.0 g	2.6 mM
DMTr	1.02 g	3 mM
DMAP	0.05 g	0.41 mM
Yield	0.6439 g	35.9 %

M.p. 78-79 ˚C, δ_H (400 MHz, CD_3CN) 8.28 (1H, s, H_1), 8.23-8.27 (2H, m, H_6), 8.03 (2H, dd, J = 6.5 & 2.9, H_3), 7.40-7.49 (6H, m, H_5, H_4 & H_{26}), 7.26-7.31 (6H, m, H_{27} & H_{21}), 7.18 (1H, t, J = 7.3, H_{28}), 6.82 (4H, d, J = 8.8, H_{22}), 6.48 (1H, d, J = 8.8, H_{14}), 4.16 (2H, t, J = 6.1, H_9), 3.90-3.99 (2H, m, H_{16} & H_{17}), 3.69 (6H, s, H_{24}), 3.05-3.19 (3H, m, H_{18} & CHOH), 2.37 (2H, t, J = 7.1, H_{12}), 1.93-2.07 (4H, m, H_{11} & H_{10}), 1.05 (3H, d, J = 6.2, H_{15}); δ_C (75 MHz, CD_3CN) 173.8 (C_{13}), 159.5 (2xC_{23}), 152.2 (C_8), 146.1 (C_{25}), 137.0 (2xC_{20}), 133.4 (2xC_2), 130.9 (4xC_{21}), 129.3 (2xC_3), 129.0 (2xC_{26}), 128.7 (2xC_{27}), 127.7 (C_{28}), 126.6 (2xC_5), 126.3 (2xC_4), 125.5 (2xC_7), 123.1 (2xC_6), 122.9 (C_1), 113.9 (4xC_{22}), 86.8 (C_{19}), 76.6 (C_9), 67.3 (C_{16}), 64.6 (C_{18}), 55.7 (2xC_{24}), 55.2 (C_{17}), 36.7 (C_{12}), 30.8 (C_{10}), 23.4 (C_{11}), 20.5 (C_{15}); *m/z* (ES+) calcd. For $C_{44}H_{45}NO_6Na$ ($M+Na^+$) 706.31, found 706.31.

(2S,3S)-1-(Bis(4-methoxyphenyl)(phenyl)methoxy)-3-hydroxybutan-2-yl 8-(anthracen-9-yloxy)octanoate – (*n* = 7D)

(JM5.4D)



JM5.3D	0.9103 g	2.15 mM
DMTr	1.0937 g	3.2 mM
DMAP	26.2 mg	0.2 mM
Yield	825.7 mg	56 %

(R_f = 0.09 in hexane/EtOAc/ TEA 40:59:1); M.p. 69-70 °C; ν_{max}/cm^{-1} 3340 (OH), 2928 (N-H), 1607 (C=O) 1507 (C=C); δ_H (300 MHz, CD_3CN) 8.23-8.27 (3H, m, H_1 & H_6), 7.96-8.00 (2H, m, H_3), 7.42-7.48 (6H, m, H_4, H_5 & H_{29}), 7.17-7.32 (7H, m, H_{24}, H_{30} & H_{31}), 6.82 (4H, d, J = 8.91, H_{25}), 6.53 (1H, d, J = 8.81, H_{17}), 4.06-4.13 (2H, m, H_9), 3.98-4.02 (2H, m, H_{19} & H_{20}), 3.69 (6H, s, H_{27}), 3.06-3.22 (3H, m, H_{21} & OH), 2.22 (2H, t, J = 7.38, H_{15}), 1.91-1.96 (2H, m, H_{10}), 1.52-1.66 (4H, m, H_{11} & H_{14}), 1.37-1.39 (4H, m, H_{12} & H_{13}), 1.08 (3H, d, J = 6.15, H_{18}); δ_C (400 MHz, CD_3CN) 172.9 (C_{16}), 158.2 (2xC_{26}), 151.0 (C_8), 144.9 (C_{28}), 135.8 (2xC_{23}), 132.1 (2xC_2), 129.8 (2xC_{30}), 128.0 (2xC_{29}), 127.7 (2xC_3), 127.4 (4xC_{24}), 126.4 (C_{31}), 125.3 (2xC_4), 124.9 (2xC_5), 124.2 (2xC_7), 121.9 (C_1), 121.5 (2xC_6), 112.7 (4xC_{23}), 85.5 (C_{22}), 75.6 (C_9), 66.0 (C_{20}), 63.3 (C_{21}), 54.5 (C_{27}), 35.7 (C_{15}), 29.9 (C_{10}), 28.7 (C_{12}), 28.7 (C_{13}), 25.5 (C_{11}), 25.4 (C_{14}), 19.4 (C_{18}). *m/z* (ES+) calcd. for $C_{47}H_{51}NO_6Na$ (M+Na$^+$) 748.3614, found 748.3612.

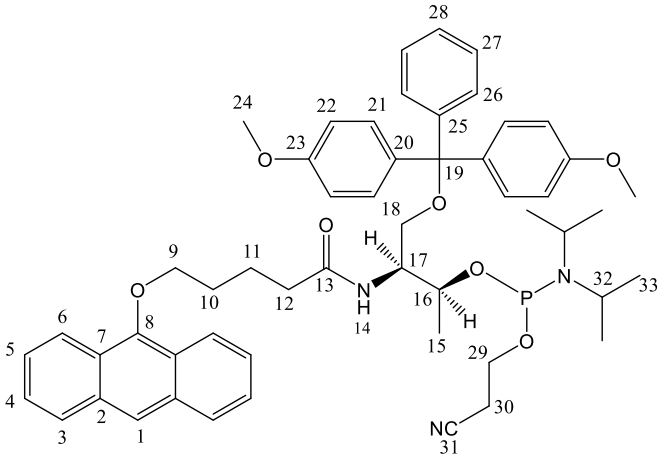
7.2.6 Synthesis of Anthracene Phosphoramidite **6**

The DMTr protected product **5** was placed in a schlenk flask with a stirrer bar and fitted with a rubber septum. The flask was evacuated and filled with argon three times and the solid was dissolved in anhydrous DCM (15 mL). DIPEA was then added to the solution and stirred in the absence of light for 15 min, followed by the drop wise addition of *2*-cyanoethoxy-N,N-diisopropylaminochlorophosphine (Phosphitylating reagent). The reaction was then left to stir for 1 hour. On completion the solution was transferred from the schlenk flask to a 25 mL round bottom flask containing a stirrer bar and benzyl alcohol (polymer bound) and was left to stir for 1 hour in the absence of light. The solution was then diluted with ethyl acetate (10 mL) and washed with a 2M Na_2CO_3 (a.q.)

solution (2 x 50 mL) and brine (1 x 50 mL) and then dried over $MgSO_4$. The resulting solution was columned through activated basic alumina (ethyl acetate/hexane/TEA, 49:50:1) and the filtrate was evaporated in *vacuo* to give a fluffy, yellow, crystalline solid.

(2S,3S)-3-(5-(Anthracen-9-yloxy)pentanamido)-4-(bis(4-methoxyphenyl)(phenyl)methoxy)butan-2-yl (2-cyanoethyl) diisopropylphosphoramidite – (*n* = 4D)

(JM1.5D)



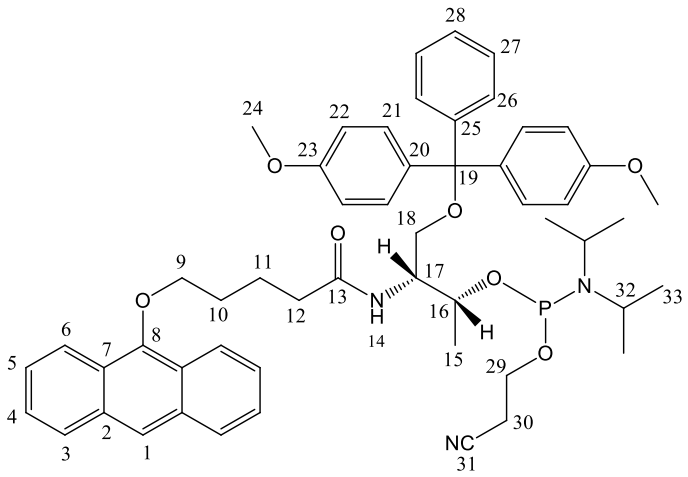
JM1.4D	0.4539 g	0.66 mM
DIPEA	0.56 mL	3.32 mM
Phosphytilating reagent	0.2 mL	0.9 mM
Yield	352 mg	59.9 %

δ_H (300 MHz, CD_3CN) 8.32 (1H, s, H_1), 8.30-8.25 (2H, m, H_6), 8.07 (2H, dd, J = 6.3 & 3.3, H_3), 7.57-7.42 (4H, m, H_4 & H_{26}), 7.39-7.26 (6H, m, H_5 & H_{21}), 7.21 (3H, td, J = 7.2 & 5.0 H_{27} & H_{28}), 6.84 (4H, ddd, J = 6.9, 5.0 & 2.6, H_{22}), 6.26 (1H, dd, J = 16.6 & 9.2, H_{14}), 4.22 (2H, dt, J = 12.1 & 6.2, H_9), 3.80-3.63 ("H, m, H_{17} & H_{16}), 3.73 (6H, s, H_{24}), 3.43-3.60 (4H, m, H_{29} & H_{32}), 3.06-3.21 (2H, m, H_{18}), 2.52 (2H, dt, J = 33.9 & 6.1, H_{30}), 2.41 (2H, t, J = 6.0, H_{12}), 2.02-

2.11 (4H, m, H_{11} & H_{10}), 1.29-1.05 (12H, m, H_{33}), 0.98 (3H, d, J = 6.8, H_{15}); δ_P (121 MHz, CD_3CN) 147.6, 147.0; *m/z* (ES+) calcd. For $C_{54}H_{64}O7N_3NaP$ ($M+Na^+$) 906.42, found 906.42.

(2R,3R)-3-(5-(Anthracen-9-yloxy)pentanamido)-4-(bis(4-methoxyphenyl)(phenyl)methoxy)butan-2-yl (2-cyanoethyl) diisopropylphosphoramidite – (*n* = 4L)

(JM1.5L)



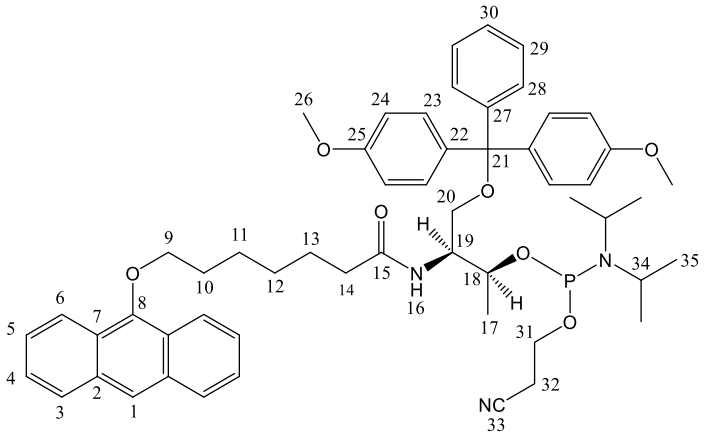
JM1.4L	0.6439 g	0.94 mM
DIPEA	0.82 mL	4.7 mM
Phosphytilating reagent	0.43 mL	1.9 mM
Yield	376.4 mg	45.3%

δ_H (300 MHz, CD_3CN) 8.32 (1H, s, H_1), 8.30-8.25 (2H, m, H_6), 8.07 (2H, dd, J = 6.3 & 3.3, H_3), 7.57-7.42 (4H, m, H_4 & H_{26}), 7.39-7.26 (6H, m, H_5 & H_{21}), 7.21 (3H, td, J = 7.2 & 5.0 H_{27} & H_{28}), 6.84 (4H, ddd, J = 6.9, 5.0 & 2.6, H_{22}), 6.26 (1H, dd, J = 16.6 & 9.2, H_{14}), 4.22 (2H, dt, J =12.1 & 6.2, H_9), 3.80-3.63 (2H, m, H_{17} & H_{16}), 3.73 (6H, s, H_{24}), 3.43-3.60 (4H, m, H_{29} and H_{32}), 3.06-3.21 (2H, m, H_{18}), 2.52 (2H, dt, J = 33.9 & 6.1, H_{30}), 2.42 (2H, t, J = 6.0, H_{12}) 2.02-2.11 (4H, m, H_{11} & H_{10}), 1.29-1.05 (12H, m, H_{33}), 0.98 (3H, d, J = 6.8, H_{15}); δ_P (121 MHz,

CD_3CN) 147.6, 147.0; *m/z* (ES+) calcd. For $C_{54}H_{64}O_7N_3NaP$ ($M+Na^+$) 906.42, Found 906.42.

(2S,3S)-3-(7-(Anthracen-9-yloxy)heptanamido)-4-(bis(4-methoxyphenyl)(phenyl)methoxy)butan-2-yl (2-cyanoethyl) diisopropylphosphoramidite – (*n* = 6D)

(JM2.1D)

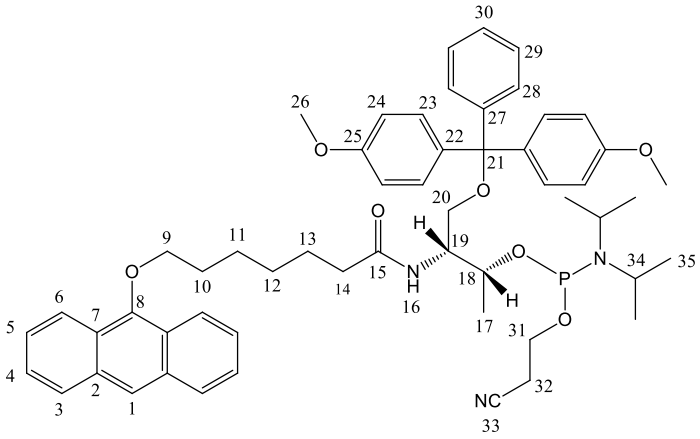


JL1	0.5 g	0.42 mM
DIPEA	0.3 mL	1.69 mM
Phosphytilating reagent	0.1 mL	0.46 mM
Yield	200.7 mg	52.4 %

δ_H (300 MHz, CD_3CN) 8.32 (1H, s, H_1), 8.26-8.31 (2H, m, H_6), 8.02-8.12 (2H, m, H_3), 7.42-7.57 (4H, m, H_4 & H_{28}), 7.26-7.38 (6H, m, H_5 & H_{23}), 7.15-7.24 (3H, m, H_{29} and H_{30}), 6.80-6.91 (4H, m, H_{24}), 6.19 (1H, dd, *J* = 15.4, & 9.2, H_{16}), 4.10-4.23 (4H, m, H_9, H_{19} & H_{18}), 3.75 (6H, s, H_{26}), 3.49-3.63 (4H, m, H_{31} & H_{34}), 3.04-3.15 (2H, m, H_{20}), 2.52 (2H, dt, *J* = 34.7 & 6.0, H_{32}), 2.28 (2H, t, *J* = 7.3 H_{14}), 1.99-2.06 (2H, m, H_{10}), 1.62-1.79 (4H, m, H_{11} & H_{13}), 1.51 (2H, dd, *J* = 15.2 & 8.1, H_{12}), 1.13 (12H, m, H_{35}), 0.99 (3H, d, *J* = 6.8, H_{17}); δ_P (121 MHz, CD_3CN) 147.6, 146.9; *m/z* (ES+) calcd. For $C_{55}H_{66}O_7N_3NaP$ (M+Na+) 934.45, found 934.45.

(2R,3R)-3-(7-(Anthracen-9-yloxy)heptanamido)-4-(bis(4-methoxyphenyl)(phenyl)methoxy)butan-2-yl (2-cyanoethyl) diisopropylphosphoramidite – (*n* = 6L)

(JM2.1L)

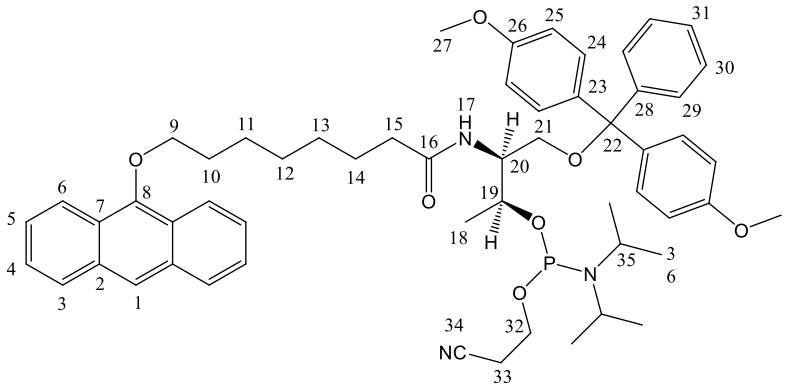


JL2	0.5 g	0.42 mM
DIPEA	0.3 mL	1.69 mM
Phosphytilating reagent	0.1 mL	0.46 mM
Yield	217.2 mg	56.7 %

δ_H (300 MHz, CD_3CN) 8.29 (1H, s, H_1), 8.23-8.28 (2H, m, H_6), 8.01-8.08 (2H, m, H_3), 7.39 7.53 (4H, m, H_4 & H_{28}), 7.25-7.33 (6H, m, H_5 & H_{23}), 7.22 (3H, m, H_{29} & H_{30}), 6.80-6.91 (4H, m, H_{24}), 6.19 (1H, dd, J = 15.4 & 9.2, H_{16}), 4.10-4.23 (4H, m, H_9, H_{19} & H_{18}), 3.75 (6H, s, H_{26}), 3.49-3.63 (4H, m, H_{31} & H_{34}), 3.12 (2H, m, H_{20}), 2.52 (2H, dt, J = 34.7 and 6.0, H_{32}), 2.28 (2H, t, J = 7.3 H_{14}), 1.92-2.06 (2H, m, H_{10}), 1.62-1.79 (4H, m, H_{11} and H_{13}), 1.51 (2H, dd, J = 15.2 and 8.1, H_{12}), 1.13 (12H, m, H_{35}), 0.99 (3H, d, J = 6.8, H_{17}); δ_P (121 MHz, CD_3CN) 147.7, 47.0; *m/z* (ES+) calcd. For $C_{55}H_{66}O_7N_3NaP$ ($M+Na^+$) 934.45, found 934.45.

(2S,3S)-1-(Bis(4-methoxyphenyl)(phenyl)methoxy)-3-(((2-cyanoethoxy)(diisopropylamino)phosphino)oxy)butan-2-yl 8-(anthracen-9-yloxy)octanoate

(JM5.5D)



JM5.4D	798 mg	1.1 mM
DIPEA	0.8 mL	4.6 mM
Phosphytilating reagent	0.4 mL	1.8 mM
Yield	833 mg	78.03 %

(R_f = 0.64 in pet. ether with 50% DCM); ν_{max}/cm^{-1} 2931 (N-H), 1607 (C=O), 1507 (C=C) δ_H (300 MHz, CD_3CN) 8.31 (1H, s, H_1), 8.28-8.30 (2H, m, H_6), 8.04-8.08 (2H, m, H_3), 7.50-7.53 (4H, m, H_4 & H_5), 7.44-7.48 (2H, m, H_{29}), 7.30-7.35 (6H, m, H_{24} & H_{30}), 7.20-7.24 (1H, m, H_{31}), 6.84-6.88 (4H, m, H_{25}), 6.19 (1H, dd, J = 15.58 & 9.12, H_{17}), 4.06-4.24 (4H, m, H_9, H_{19} & H_{20}), 3.75 (6H, s, H_{27}), 3.48-3.58 (4H, m, H_{32} & H_{35}), 3.07-3.19 (2H, m, H_{21}), 2.55 (2H, dt, J = 35.11 & 5.95, H_{33}), 2.23-2.29 (2H, m, H_{15}), 1.96-2.00 (2H, m, H_{10}), 1.59-1.70 (4H, m, H_{11} & H_{14}), 1.44-1.46 (4H, m, H_{12} & H_{13}), 1.09-1.18 (12H, m, H_{36}), 0.99 (3H, d, J = 6.75, H_{18}); δ_C (400 MHz, CD_3CN) 172.3 (C_{16}), 158.3 (2xC_{26}), 151.0 (C_8), 144.9 (C_{28}), 135.8 (2xC_{23}), 132.1 (2xC_2), 129.7 (2xC_{30}), 128.0 (2xC_{29}), 127.7 (2xC_3), 127.4 (4xC_{24}), 126.4 (C_{31}), 125.3 (2xC_4), 124.9 (2xC_5), 124.2 (2xC_7), 121.9 (C_1), 121.5 (2xC_6), 112.6 (4xC_{25}), 112.3 (C_{34}), 85.41 (C_{22}), 75.6 (C_9), 68.7 (C_{19}), 63.1 (C_{21}), 62.8 (C_{20}), 58.1 (C_{32}), 54.5 (2xC_{27}), 53.4 (2xC_{35}), 35.7 (C_{15}), 29.9 (C_{10}), 28.7 (C_{12}), 28.6 (C_{13}), 25.4 (C_{11}), 25.3 (C_{14}), 23.6 (4xC_{36}), 19.6

(C_{33}), 18.2 (C_{18}); δ_P (300 MHz, CD_3CN) 147.6, 146.9; *m/z* (ES+) calcd. for $C_{56}H_{68}N_3O_7NaP$ ($M+Na^+$) 948.4693, found 948.4679.

7.3 Data Processing.

Unless otherwise stated all T_m and fluorescent data is normalised in order to ease comparison between the various samples, the following section describes how the normalisation was performed.

For all T_m melting curves the initial data point is subtracted from the entire data set for each sample, therefore each transition starts at 0 and the graph displays the difference in absorbance (ΔA) rather than actual absorbance recorded. To normalise the remaining data the each data point is multiplied by a common factor (Calculated from the ratio between the last data points) so that the transition of each melting curve ends at the same point. Figure 1 displays an example Excel spreadsheet and the formulas that can be used to normalise sample B to sample A.

	A	B	C	D
1	Temperature (°C)	Sample A	Sample B	Sample B (normalised)
2	15	0.01	0.01	= C2 * C$72
3	16	0.02	0.02	= C3 * C$72
…	…	…	…	…
70	84	0.14	0.09	= C70 * C$72
71	85	0.15	0.09	= C71 * C$72
72			**= B71 / C71**	

Figure 1. An Excel spreadsheet of an example data set where sample B has been normalised against sample using a common factor calculated using the formula highlighted in red.

For fluorescence spectra the same set of calculations can be used to normalise the data between different samples, except in this case the common factor is calculated from the data points of the emission maxima (λ_{max}). Therefore all data sets are normalised so that they meet at the point where there is maximum fluorescent emission, which for anthracene is typically 430 nm.

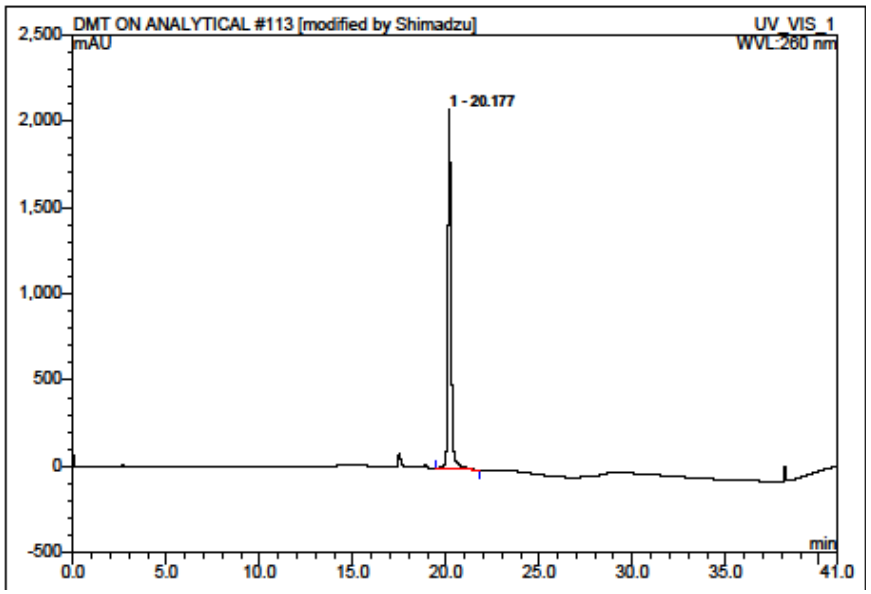
Appendix

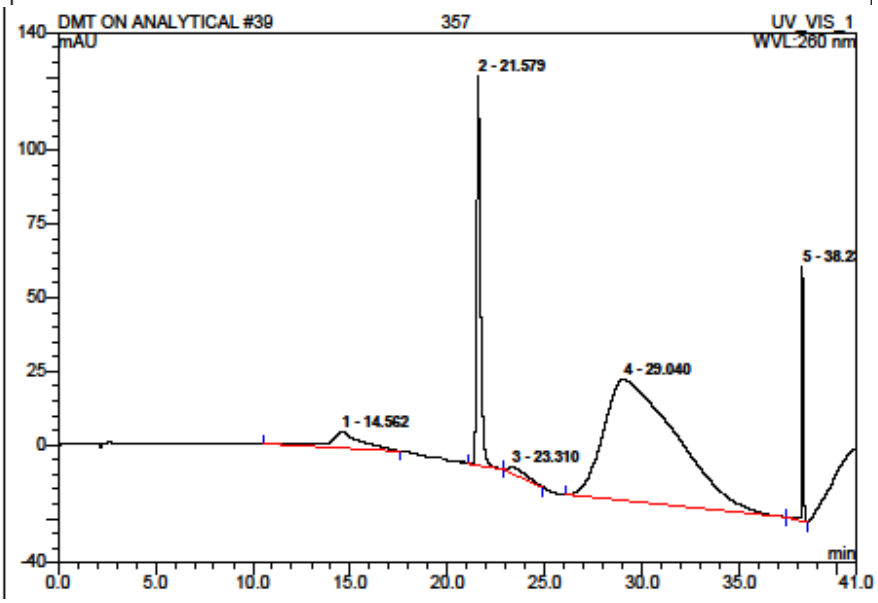
Sequence name	ID #	Sequence	Linker Type	Predicted m/z	Obs. m/z	Ret. Time	E (M^{-1})	Conc (μM)	Yield (%)
S1	274	TGGACT**X**T**X**TCAATG	4D	4851.4	4851.0	53.4	233600	149.0	22.3
S2	339	TGGACTC**X**C**X**CAATG	4D	4821.4	4821.0	19.3	231800	177.0	26.5
S3	327	TGGACTC**X**C**X**CAATG	4L	4821.4	4821.5	19.2	231800	150.3	22.5
S4	355	TGGACTC**X**C**X**CAATG	6D	4877.6	4875.0	21.2	231800	78.4	11.7
S5	348	TGGACTC**X**C**X**CAATG	6L	4877.6	4874.4	21.1	231800	107.2	16.1
S6	341	TGGACT**X**T**X**TCAATG	4D	4851.4	4850.5	19.6	233600	392.2	58.8
S7	329	TGGACT**X**T**X**TCAATG	4L	4851.4	4851.0	19.1	233600	469.1	70.3
S8	357	TGGACT**X**T**X**TCAATG	6D	4907.6	4909.0	20.8	233600	305.1	45.7
S9	350	TGGACT**X**T**X**TCAATG	6L	4907.6	4905.0	21.6	233600	140.8	21.1
S10	340	TGGAC**X**C**X**CTCAATG	4D	4821.4	4821.0	19.1	231800	392.2	58.8
S11	328	TGGAC**X**C**X**CTCAATG	4L	4821.4	4821.5	18.7	231800	164.4	24.6
S12	356	TGGAC**X**C**X**CTCAATG	6D	4877.6	4875.3	21.1	231800	254.0	38.0
S13	349	TGGAC**X**C**X**CTCAATG	6L	4877.6	4874.4	21.0	231800	245.7	36.8
S14	357	TGGACT**X**T**X**TCAATG	6D	4907.6	4909.0	21.9	233600	305.1	45.7
S15	475	TGGAC**X**CTC**X**CAATG	6D	4877.6	4877.0	20.1	231800	251.4	37.6
S16	478	TGGA**X**TCTCT**X**AATG	6D	4907.6	4907.0	20.2	235200	218.9	32.6
S14-A	514	TGGACT**X**T**Y**TCAATG	6D	4574.2	4575.0	17.6	171100	141.9	21.2
S15-A	481	TGGAC**X**CTC**Y**CAATG	6D	4544.2	4545.0	17.5	176800	156.9	23.5
S16-A	517	TGGA**X**TCTCT**Y**AATG	6D	4574.2	4575.0	17.6	182200	209.7	31.4
S14-B	515	TGGACT**Y**T**X**TCAATG	6D	4574.2	4575.0	17.3	171100	147.8	22.1
S15-B	482	TGGAC**Y**CTC**X**CAATG	6D	4544.2	4545.0	17.4	176800	228.9	34.3
S16-B	516	TGGA**Y**TCTCT**X**AATG	6D	4574.2	4576.0	17.6	182200	102.2	15.3
S14-C	518	TGGACT**Y**T**Y**TCAATG	-	4240.8	4242.0	28.2	116100	223.3	33.4

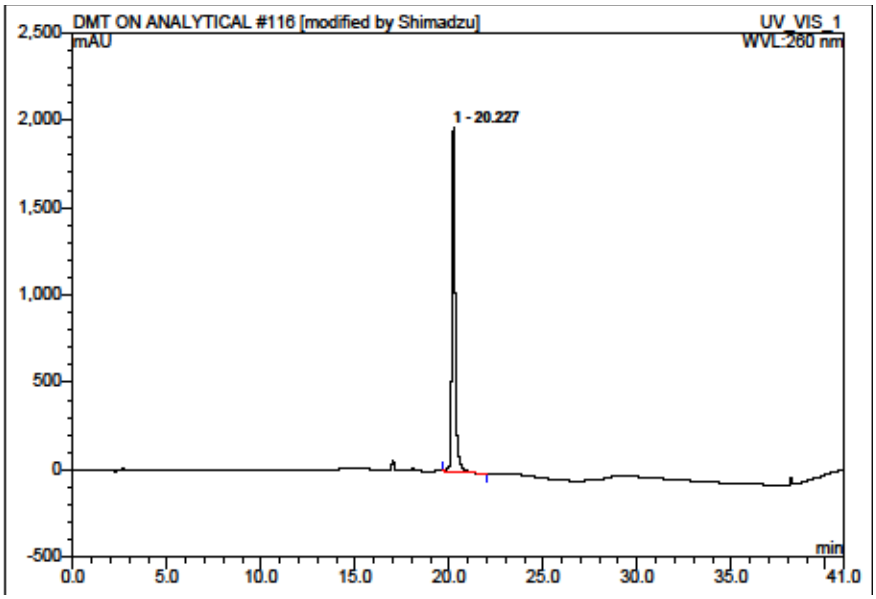
S15-C	480	TGGACYCTCYCAATG	-	4210.7	4212.0	25.8	121800	228.2	34.2
S16-C	499	TGGAYTCTCTYAATG	-	4240.8	4242.0	31.3	127200	255.6	38.3
S17	532	XTGGACTCTCTCAATGX	7D	5593.6	5594.0	23.6	247588	208.9	31.3
S18	533	XTTTTGGACTCTCTCAATGTTTX	7D	7418.8	7418.4	23.4	297988	144.5	21.6
S19	534	XTTTTTGGACTCTCTCAATGTTTTX	7D	8027.2	8025.2	22.9	314788	98.1	14.7
S20	535	XGCGACTGGACTCTCTCAATGGTCGCX	7D	8684.6	8688.0	21.2	342088	27.9	4.2
S21	559	GCGACTGGACTCTCTCAATGGTCGC	-	7634.0	7637.0	47.8	239200	409.4	61.3
S22	642	TGGACTCCCTCAATGCy5	-	5061.6	5068.0	22.3	137100	303.2	45.4
S23	643	TGGACCCTCCCAATGCy5	-	5046.6	5053.0	21.9	136200	151.7	22.7
S24	338	GGTTGGTGXGGXTGG	4D	5004.4	5004.5	18.4	236500	214.5	32.1
S25	326	GGTTGGTGXGGXTGG	4L	5004.4	5004.5	18.0	236500	151.3	22.7
S26	345	GGTTGGTGXGGXTGG	6D	5060.5	5060.0	20.5	236500	118.7	17.8
S27	333	GGTTGGTGXGGXTGG	6L	5060.5	5061.0	19.8	236500	204.9	30.7
S28	337	GGTTGGXGXGGTTGG	4D	5004.4	5004.5	18.8	235900	160.9	24.1
S29	325	GGTTGGXGXGGTTGG	4L	5004.4	5004.5	18.7	235900	146.9	22.0
S30	344	GGTTGGXGXGGTTGG	6D	5060.5	5060.5	20.9	235900	87.6	13.1
S31	332	GGTTGGXGXGGTTGG	6L	5060.5	5061.0	18.7	235900	141.3	21.2
S32	336	GGTXGGTGTGGTXGG	4D	5004.4	5005.5	18.0	237100	169.5	25.4
S33	324	GGTXGGTGTGGTXGG	4L	5004.4	5004.0	17.6	237100	163.4	24.5
S34	343	GGTXGGTGTGGTXGG	6D	5060.5	5060.0	20.0	237100	183.0	27.4
S35	331	GGTXGGTGTGGTXGG	6L	5060.5	5061.0	19.8	237100	162.8	24.4
S36	335	XTGGTTGGTGTGGTTGGTX	4D	6221.2	6221.5	20.7	269300	155.7	23.3
S37	323	XTGGTTGGTGTGGTTGGTX	4L	6221.2	6220.0	20.6	269300	266.4	39.9
S38	342	XTGGTTGGTGTGGTTGGTX	6D	6277.3	6276.5	22.8	269300	137.2	20.6
S39	330	XTGGTTGGTGTGGTTGGTX	6L	6277.3	6277.0	22.7	269300	205.7	30.8
S40	288	GGTTGGTGTGGTTGG	-	4726	4726.0	26.9	143300	297.6	44.6
S41	719	TGGACTCXCTCAATG	7D	4724.2	4724.0	18.1	182100	233.4	34.9

S42	816	TGGACTC**X**CTCAATG	4D	4682.2	4679.9	16.1	182100	133.6	20.0
S0	284	TGGACTCTCTCAATG	-	4543	4542.0	45.4	138000	371.1	55.6
T0	442	CATTGAGAGAGTCCA	-	4601	4600.0	40.0	153900	432.6	64.8
T1	648	CATTGAGAGAGTCCA**Cy3**	-	5108	5115.0	21.4	153900	178.6	26.8
T2	720	CATTGAG**X**GAGTCCA	7D	4773.3	4773.0	18.4	194100	264.2	39.6
T3	817	CATTGAG**X**GAGTCCA	4D	4726.2	4728.6	16.0	194100	159.65	23.9

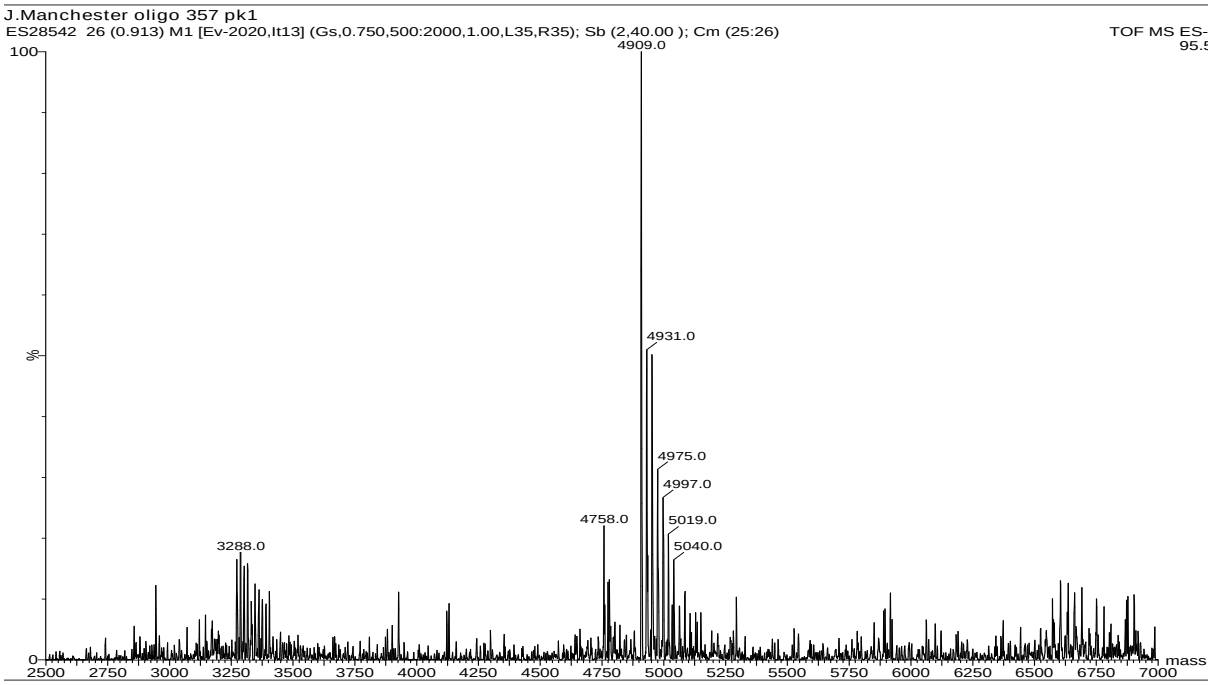
S14-S16 Analytical HPLC data

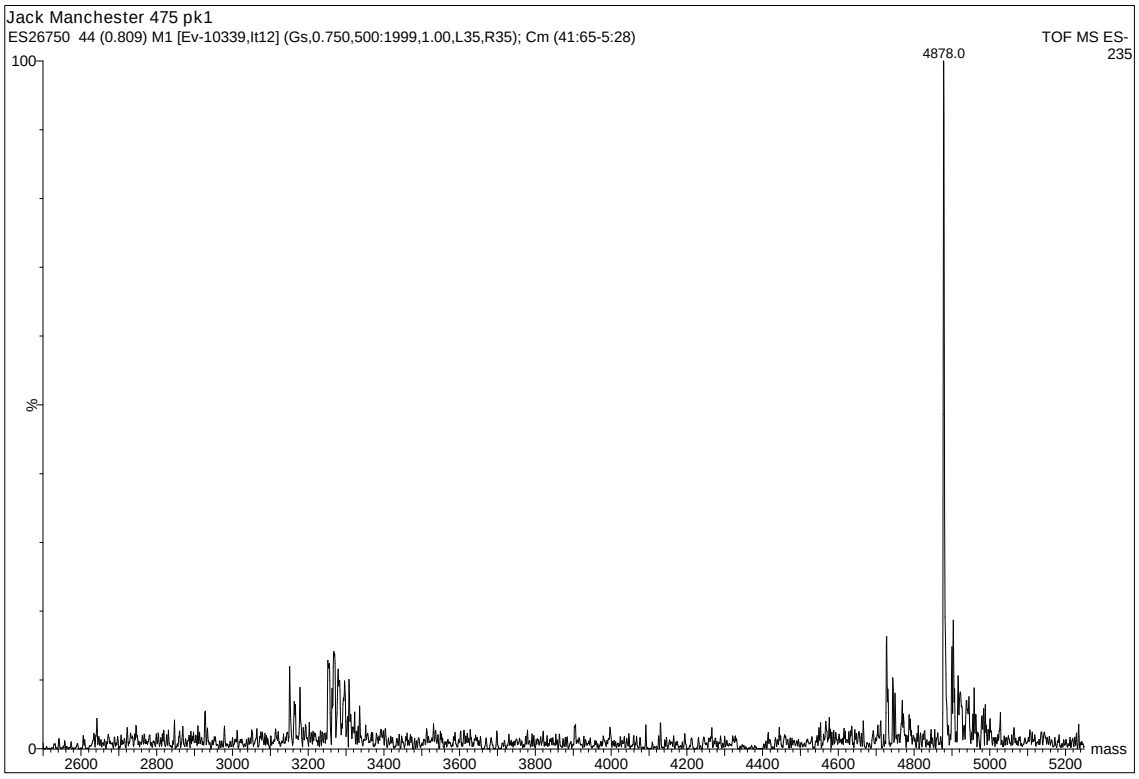


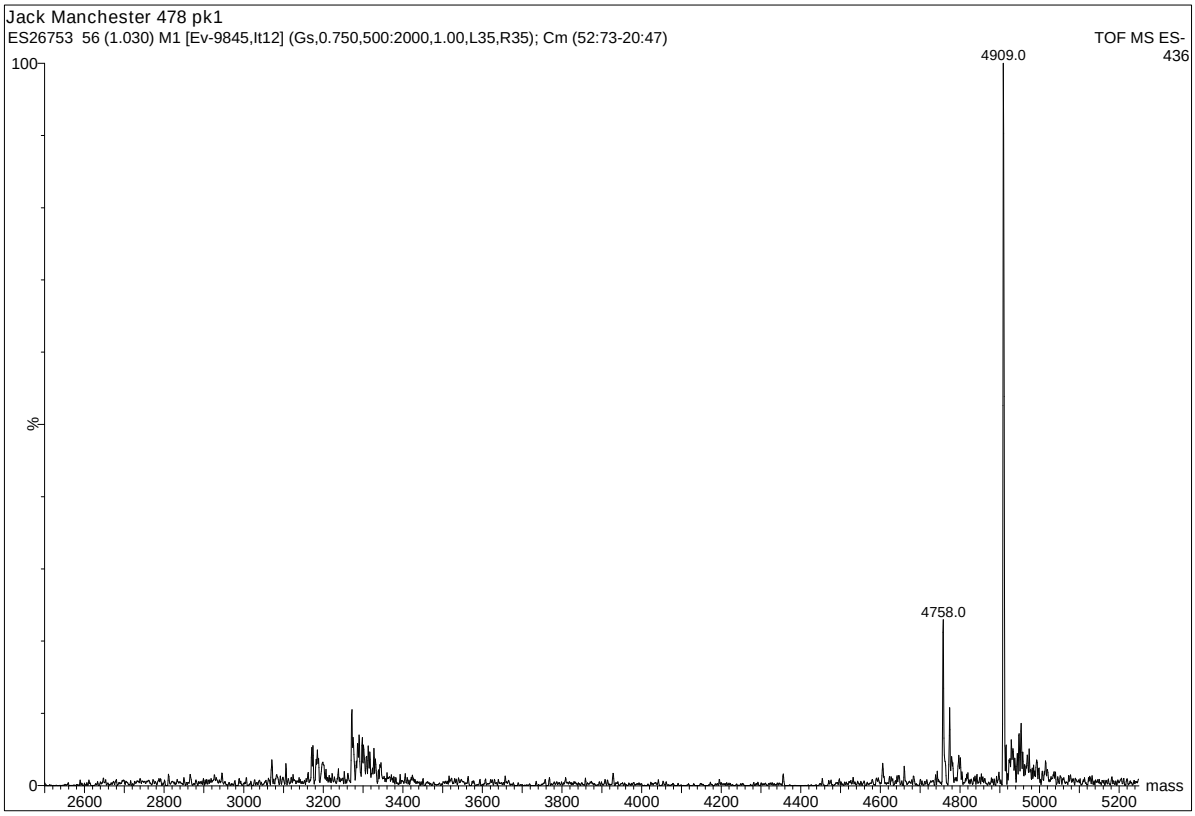




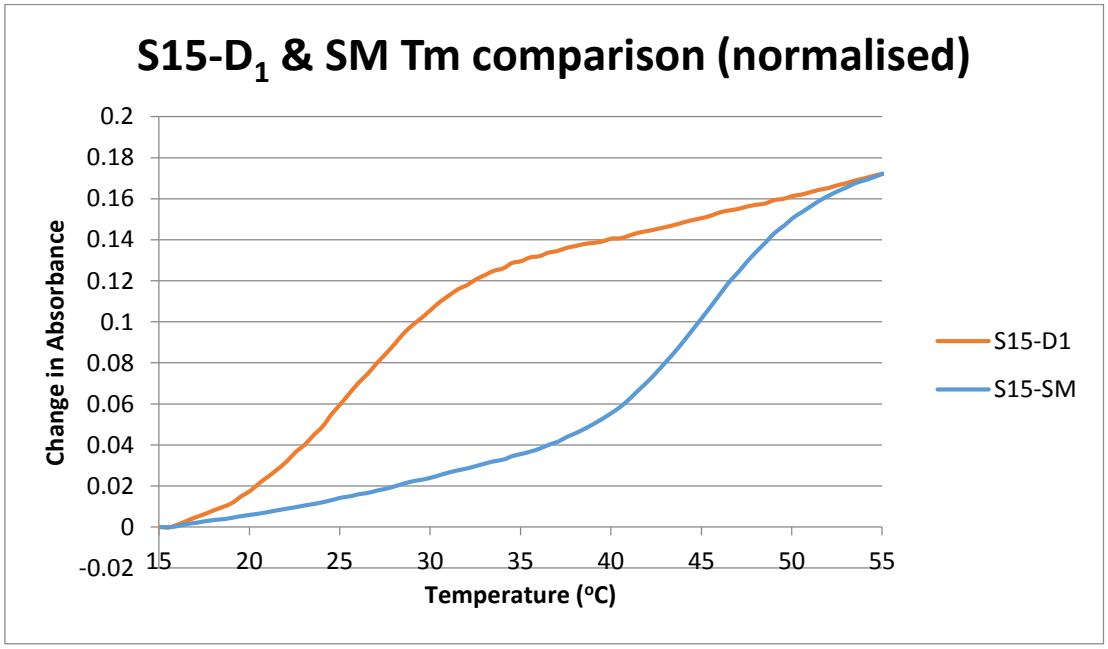
S14-S16 Photodimer mass spectrometric analysis



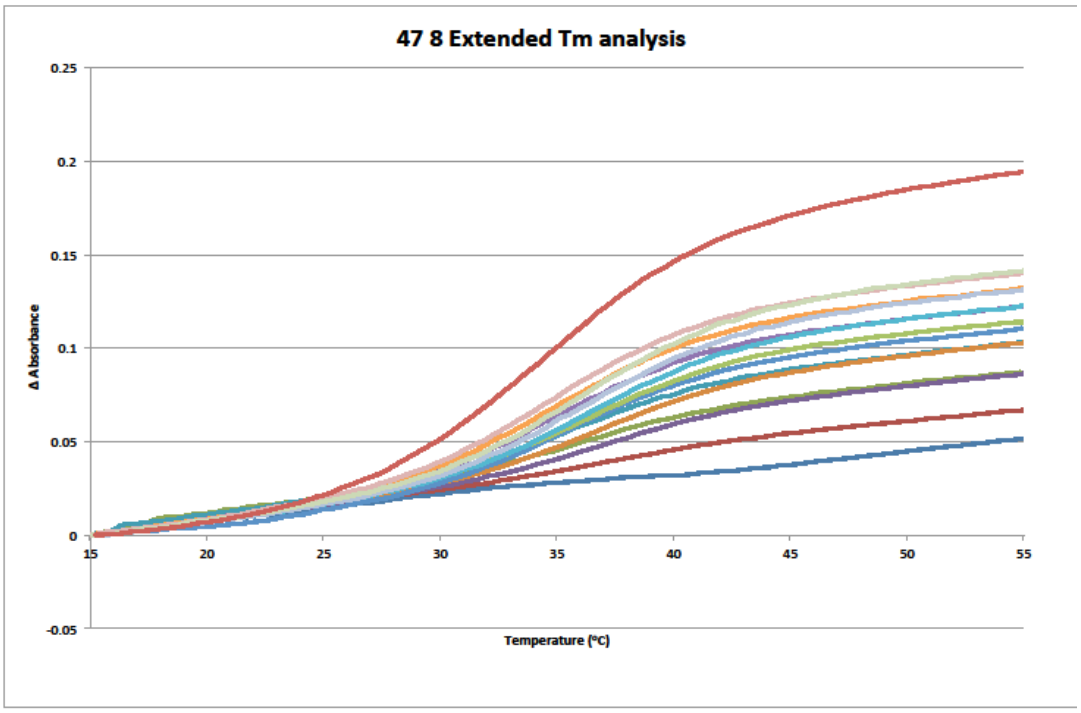




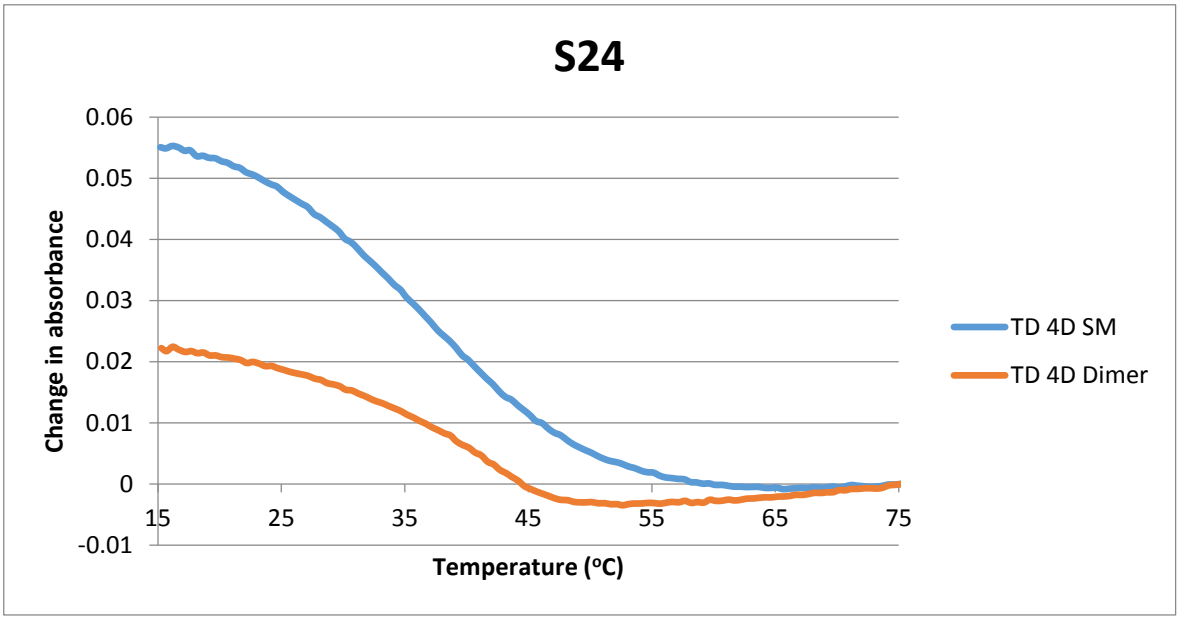
S15-D$_1$ T_m data

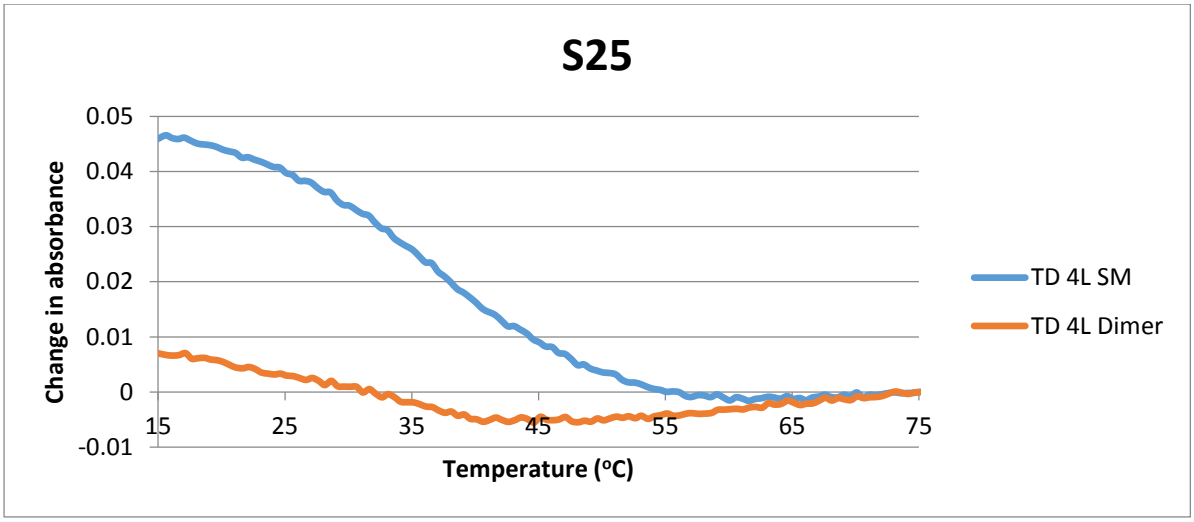


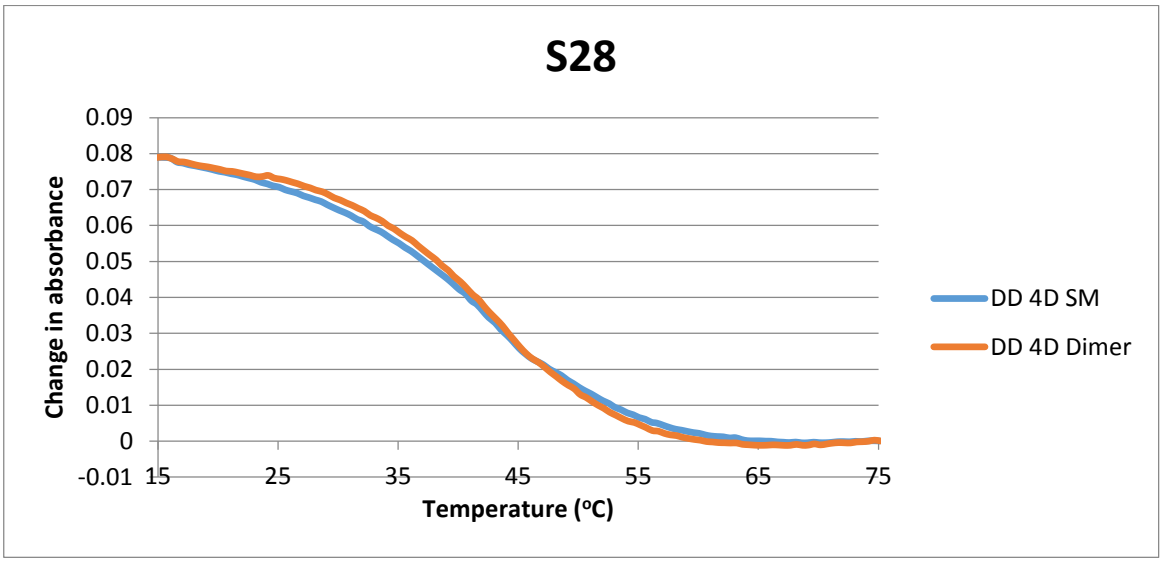
S16 Thermal Reversion T_m data

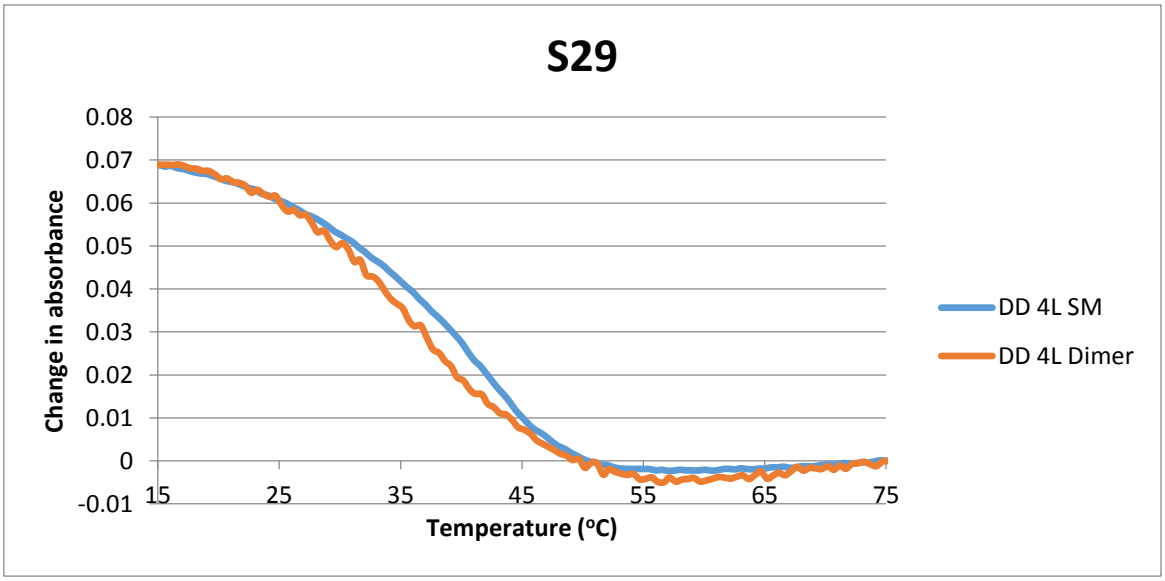


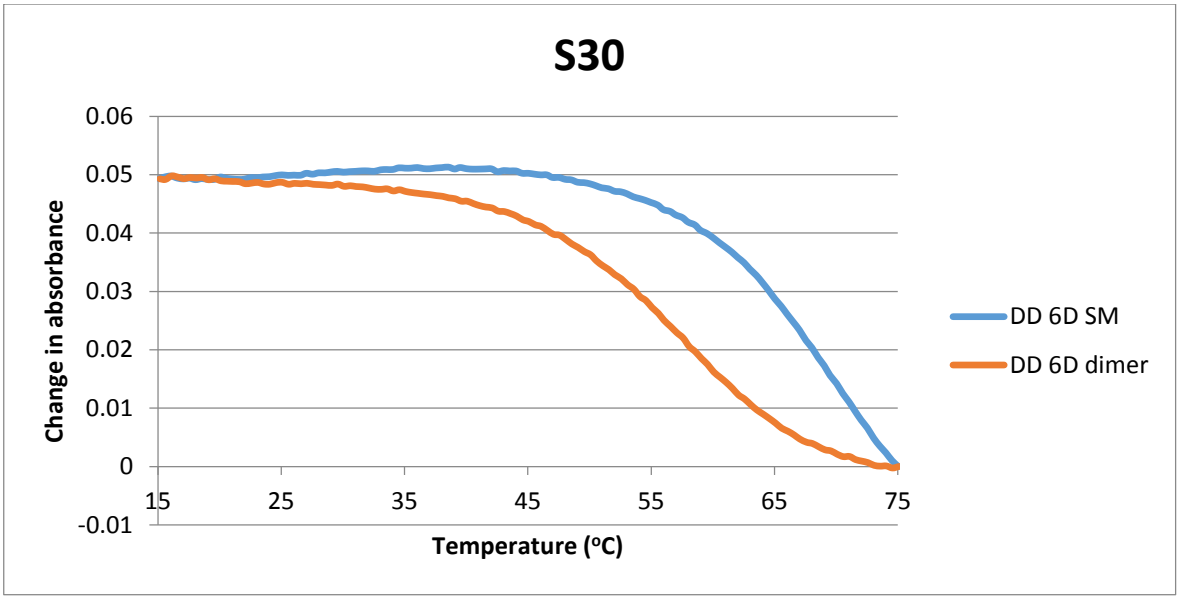
G-Quadruplex T_m data

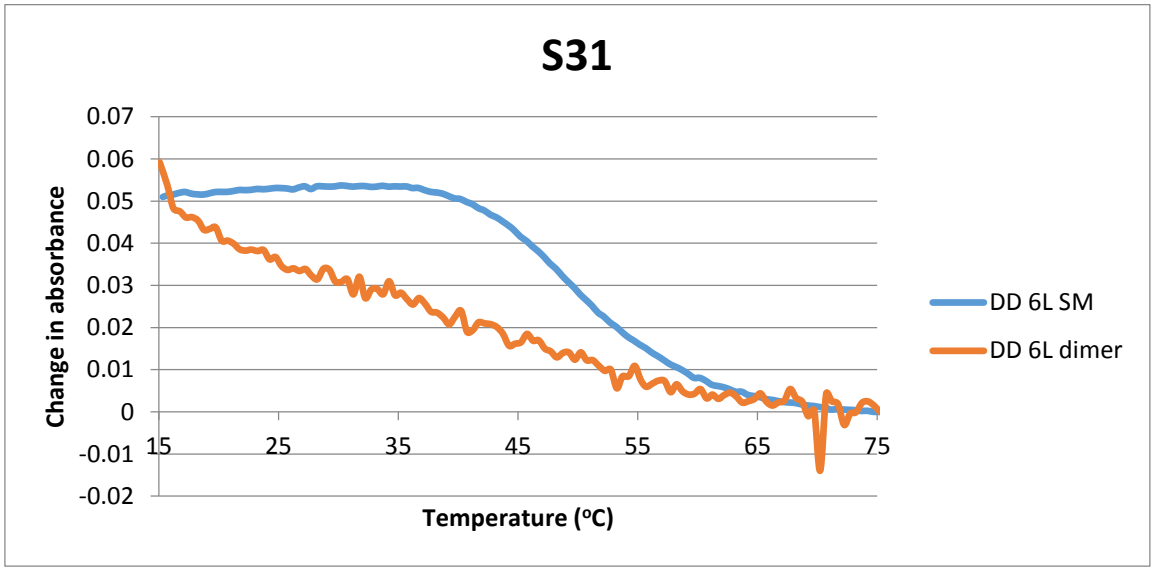


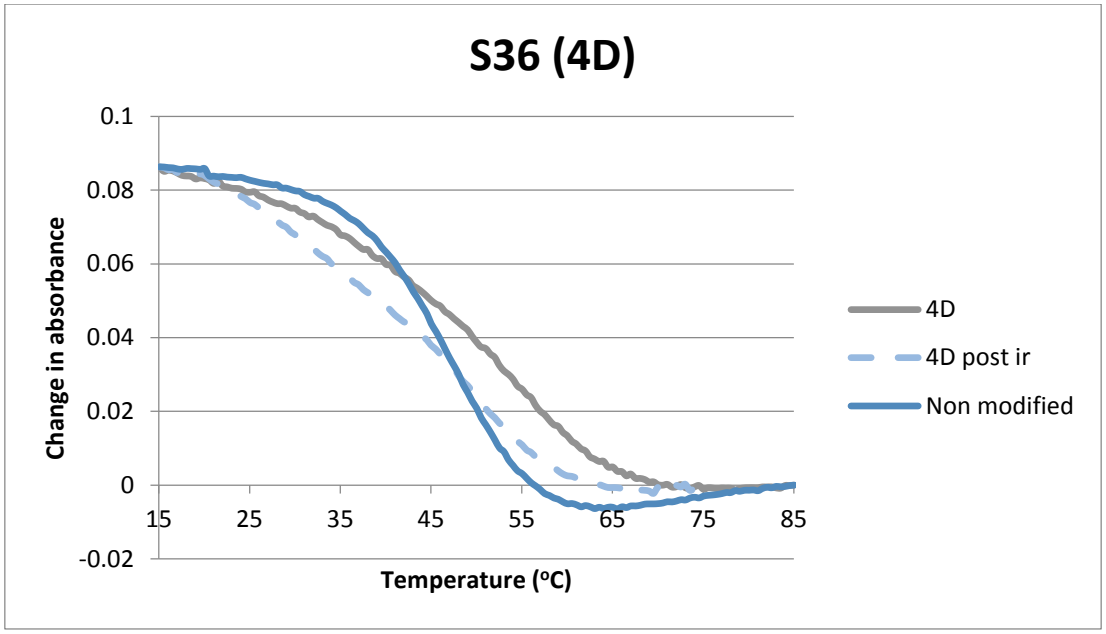

S25
0.05
0.04
0.03
0.02
0.01
0
-0.01
Change in absorbance
15
25
35
45
55
65
75
Temperature (°C)
TD 4L SM
TD 4L Dimer

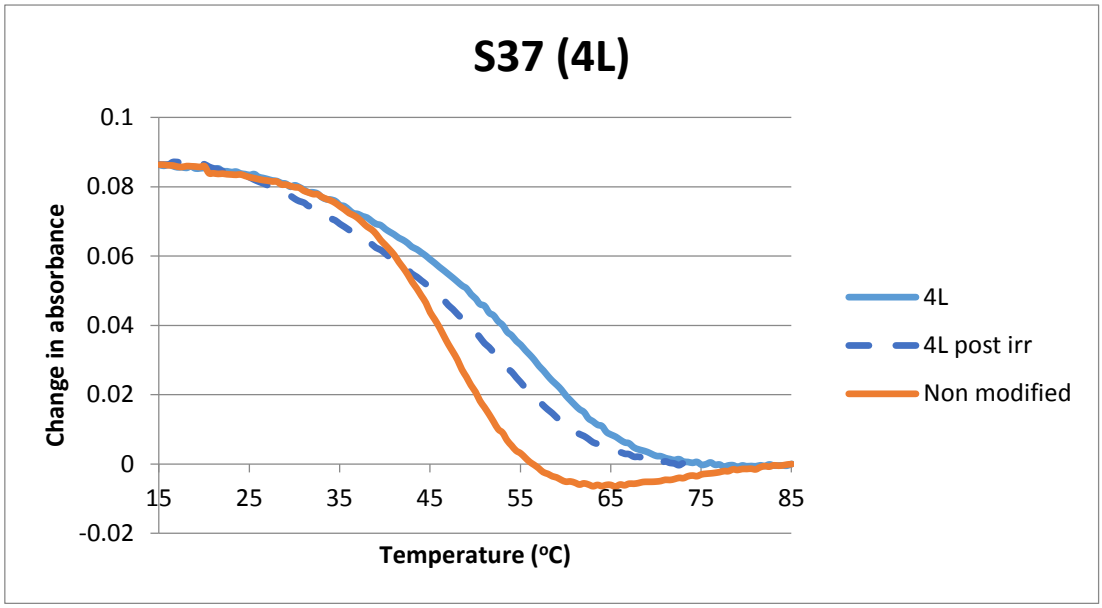

S28
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
-0.01
Change in absorbance
15
25
35
45
55
65
75
Temperature (°C)
DD 4D SM
DD 4D Dimer

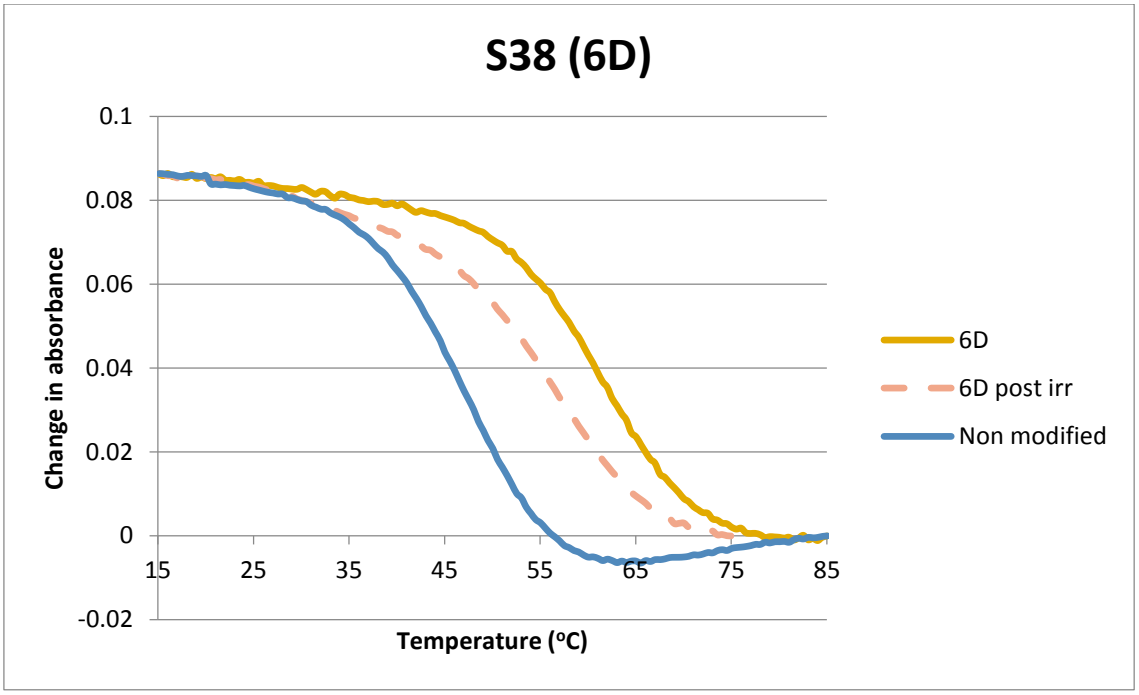

S29
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
-0.01
Change in absorbance
15
25
35
45
55
65
75
Temperature (°C)
DD 4L SM
DD 4L Dimer

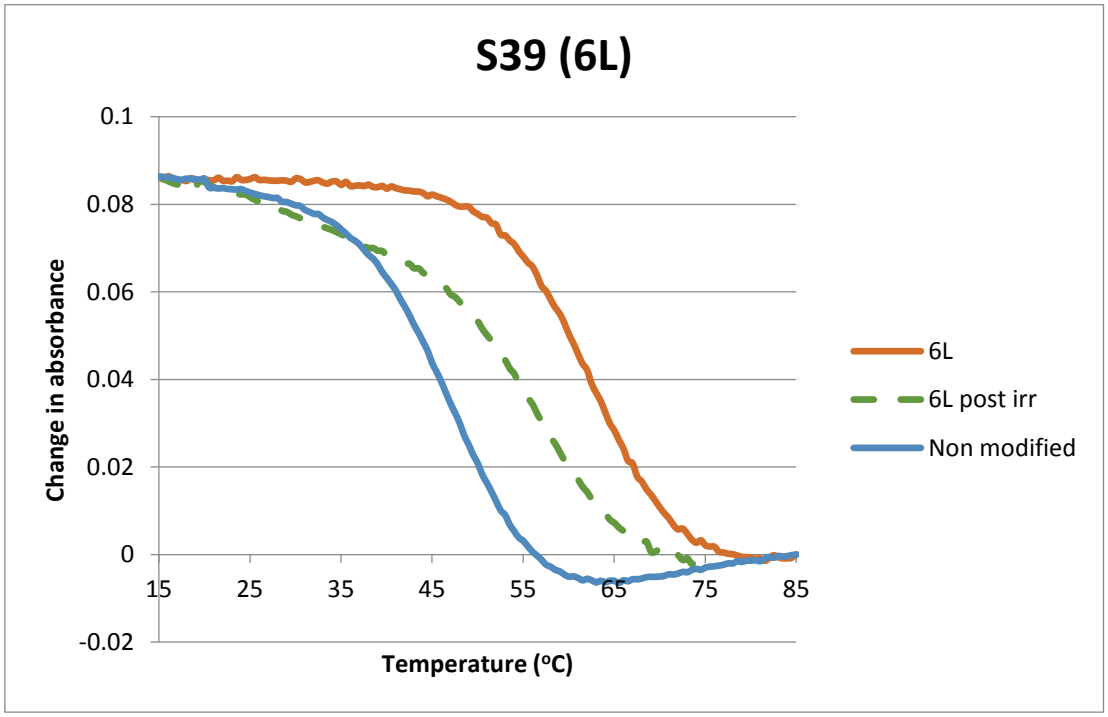

S30
0.06
0.05
0.04
0.03
0.02
0.01
0
-0.01
Change in absorbance
15
25
35
45
55
65
75
Temperature (°C)
DD 6D SM
DD 6D dimer


S31
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
-0.01
-0.02
Change in absorbance
15
25
35
45
55
65
75
Temperature (°C)
DD 6L SM
DD 6L dimer

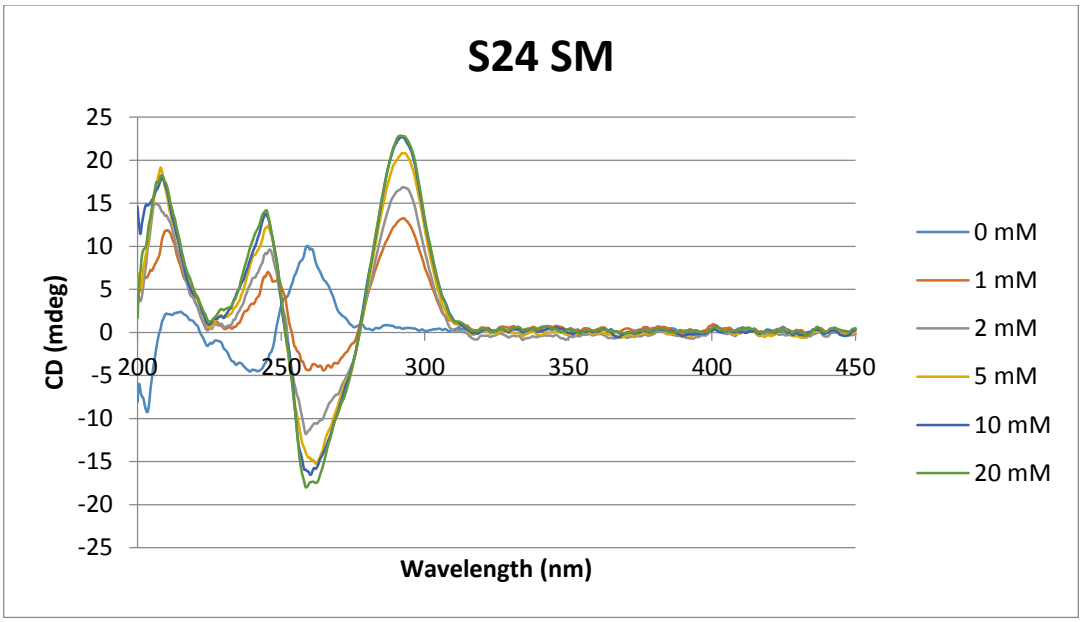

S36 (4D)
0.1
0.08
0.06
0.04
0.02
0
-0.02
Change in absorbance
15
25
35
45
55
65
75
85
Temperature (°C)
4D
4D post ir
Non modified

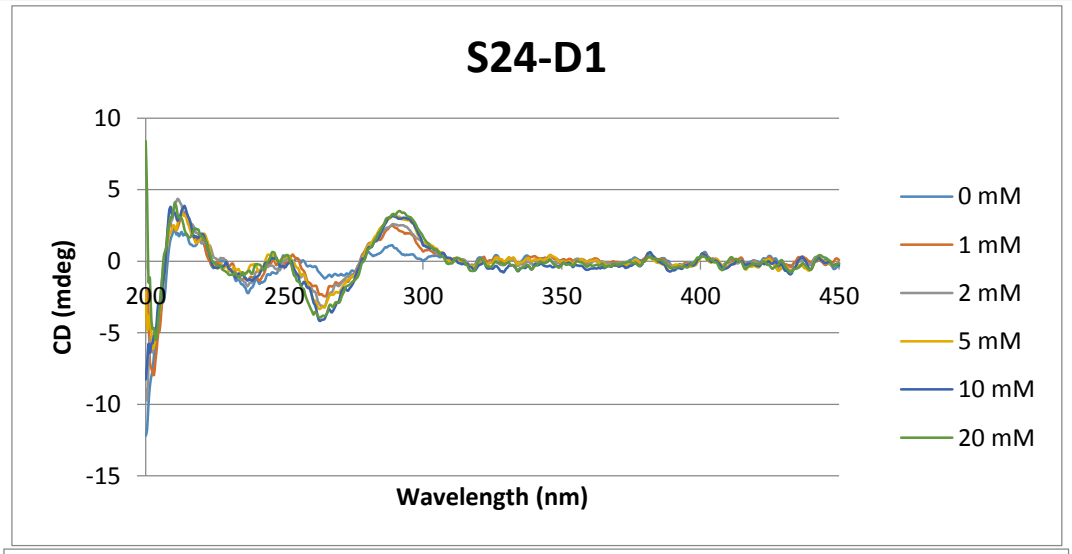

S37 (4L)
0.1
0.08
0.06
0.04
0.02
0
-0.02
Change in absorbance
15
25
35
45
55
65
75
85
Temperature (°C)
4L
4L post irr
Non modified

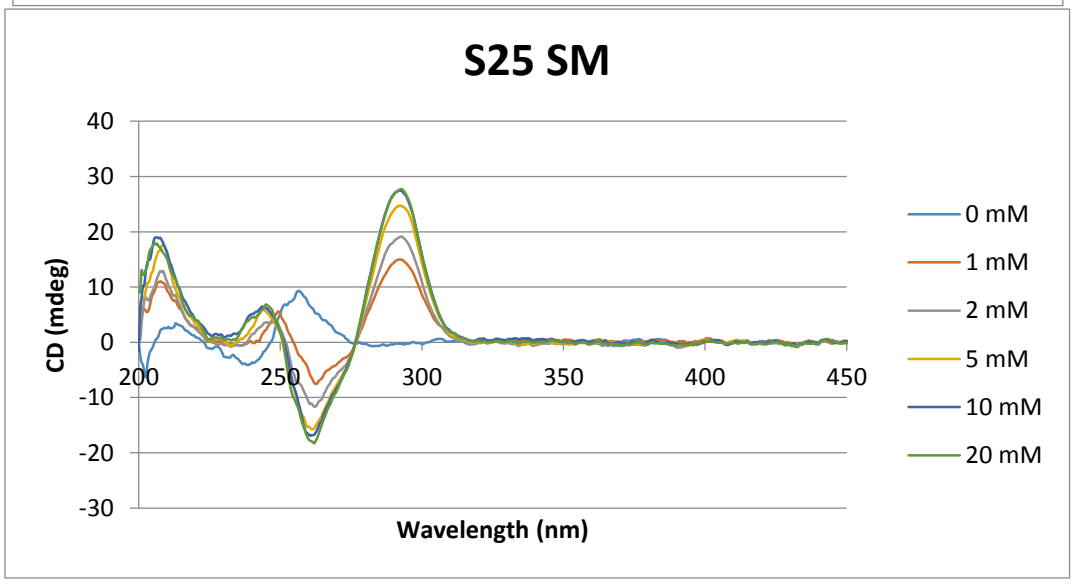

S38 (6D)
0.1
0.08
0.06
0.04
0.02
0
-0.02
Change in absorbance
15
25
35
45
55
65
75
85
Temperature (°C)
6D
6D post irr
Non modified

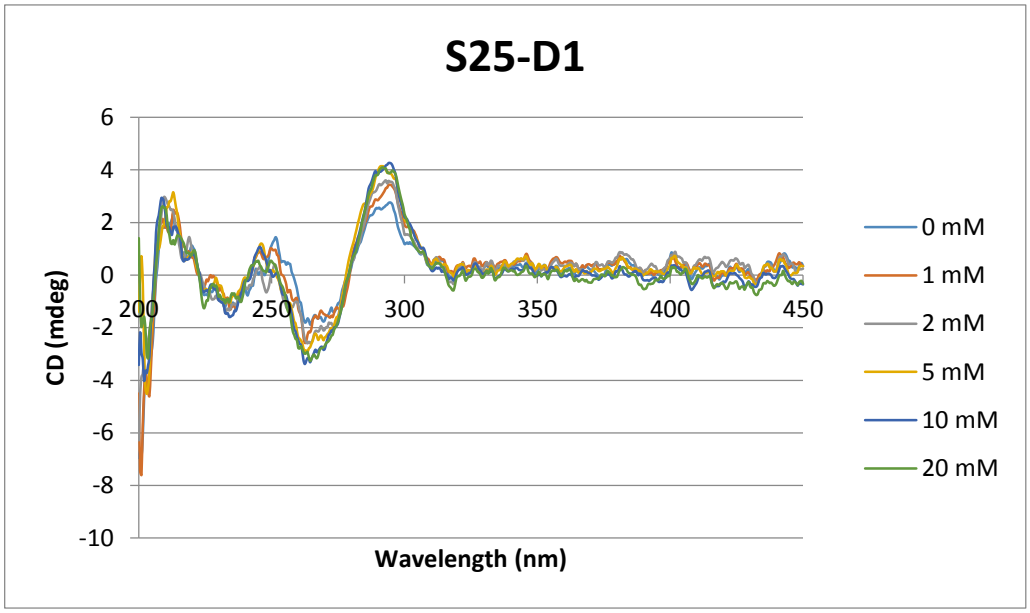

S39 (6L)
0.1
0.08
0.06
0.04
0.02
0
-0.02
Change in absorbance
15
25
35
45
55
65
75
85
Temperature (°C)
6L
6L post irr
Non modified

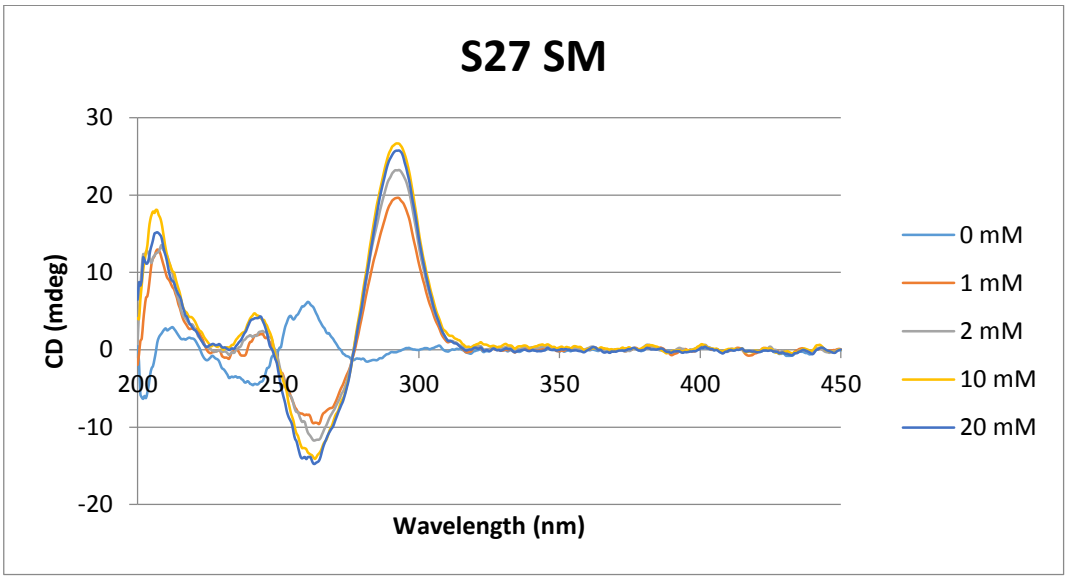
G-Quadruplex CD data

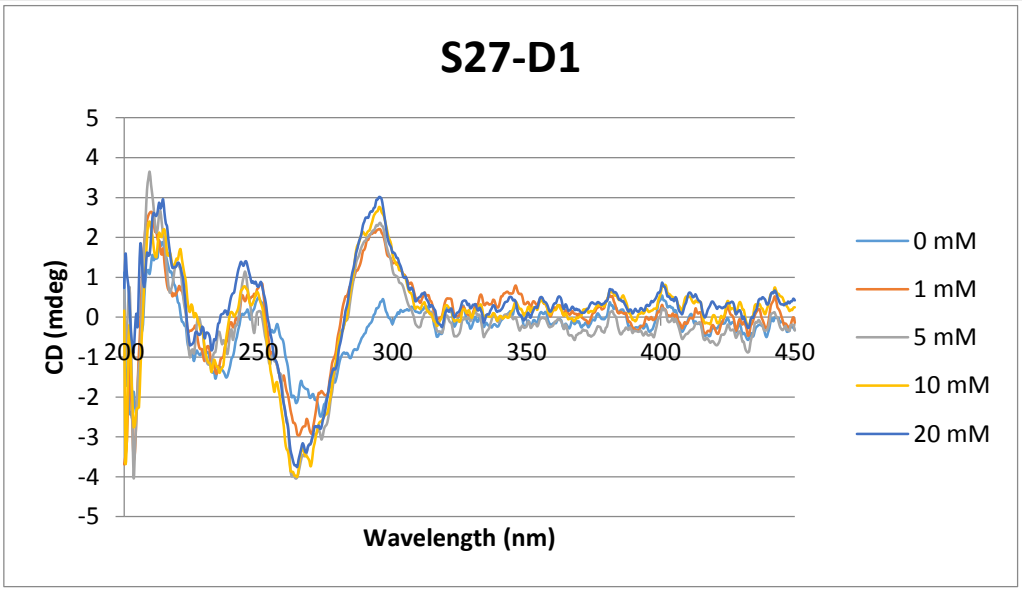


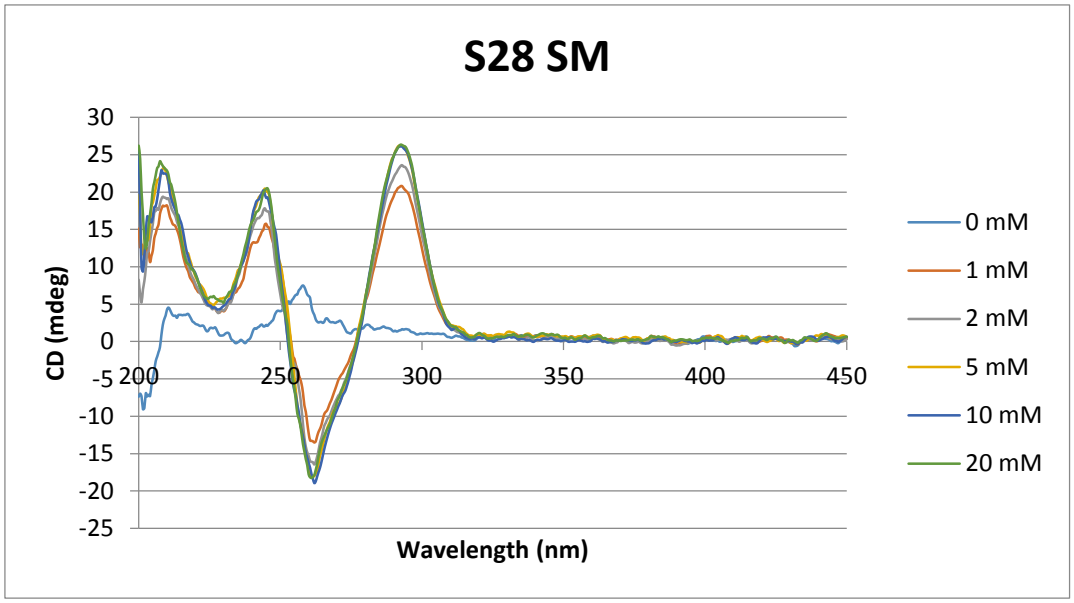


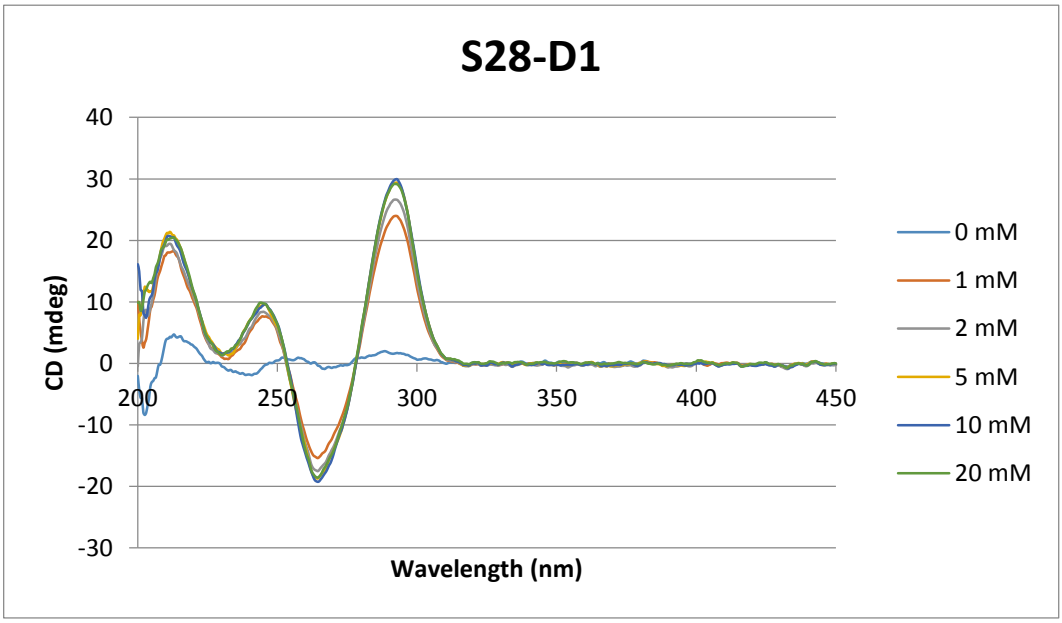


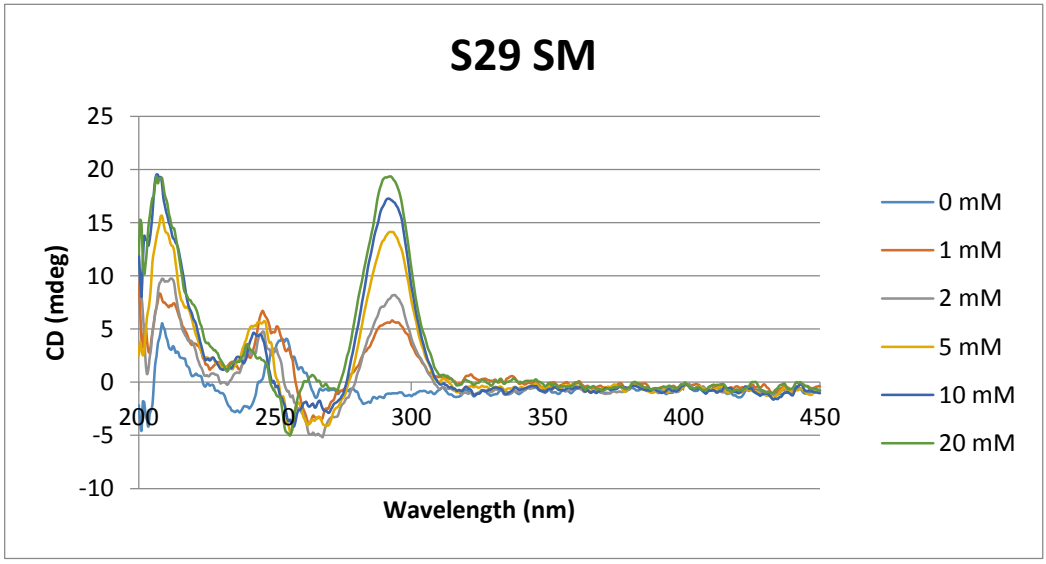

S25-D1
6
4
2
0
-2
-4
-6
-8
-10
200
250
300
350
400
450
CD (mdeg)
Wavelength (nm)
0 mM
1 mM
2 mM
5 mM
10 mM
20 mM

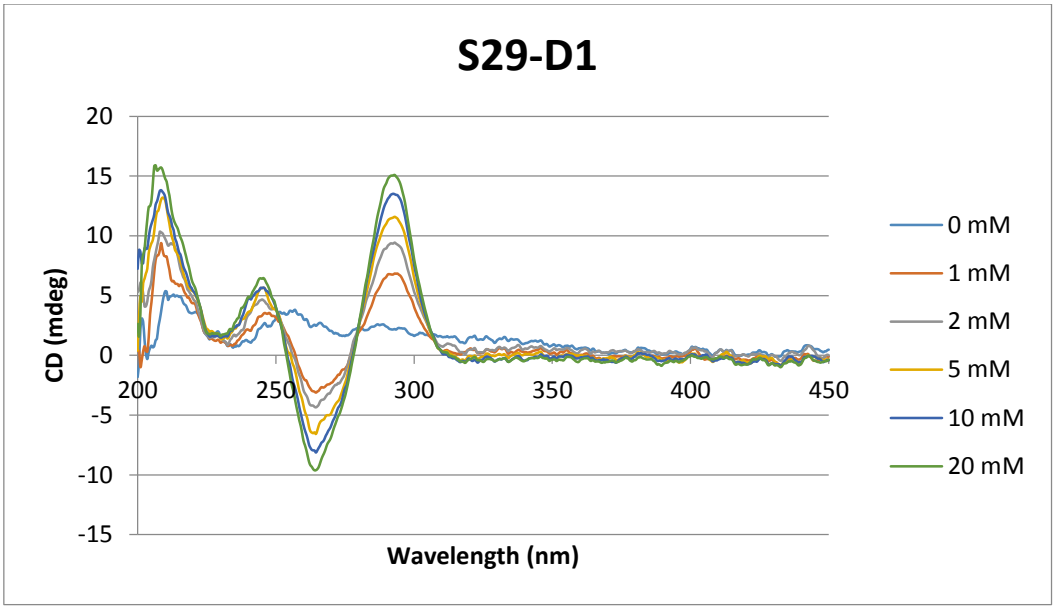

S27 SM
30
20
10
0
-10
-20
200
250
300
350
400
450
CD (mdeg)
Wavelength (nm)
0 mM
1 mM
2 mM
10 mM
20 mM


S27-D1
5
4
3
2
1
0
-1
-2
-3
-4
-5
200
250
300
350
400
450
CD (mdeg)
Wavelength (nm)
0 mM
1 mM
5 mM
10 mM
20 mM


S28 SM
CD (mdeg)
Wavelength (nm)
0 mM
1 mM
2 mM
5 mM
10 mM
20 mM


S28-D1
CD (mdeg)
Wavelength (nm)
0 mM
1 mM
2 mM
5 mM
10 mM
20 mM


S29 SM
CD (mdeg)
Wavelength (nm)
0 mM
1 mM
2 mM
5 mM
10 mM
20 mM



Stem-Loop competing strand irradiation data